华章IT

中国开源软件推进联盟PostgreSQL分会特聘专家撰写，国内多位开源数据库专家鼎力推荐。

基于PostgreSQL 10 编写，重点介绍SQL高级特性、并行查询、分区表、物理复制、逻辑复制、备份恢复、高可用、性能优化、PostGIS等，涵盖大量实战用例。

数据库技术丛书

PostgreSQL实战

谭峰 张文升 编著

机械工业出版社
China Machine Press

[illegible]

PostgreSQL[illegible]

[illegible] [illegible] [illegible]

ISBN[illegible]978-7-111-60346-7

[illegible]2018[illegible]

[illegible]

[illegible]+ 86-10-68995265

[illegible]service@bbbvip.com

[illegible]www.hzmedia.com.cn

[illegible] @[illegible]

[illegible] [illegible]hzebook[illegible]

□□

[illegible]

[illegible]Postgres[illegible]PG[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PG[illegible]

[illegible]PG[illegible]2011[illegible]6[illegible]PG[illegible]PostgreSQL[illegible]“Postgres[illegible]”[illegible]

[illegible]2[illegible]PostgreSQL[illegible]PostgreSQL[illegible]

[illegible]2011[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]7[illegible]

[illegible]Leader[illegible]Bruce Momjian[illegible]Oleg Bartunov[illegible]Simon Riggs[illegible]Magnus Hagander

Postgres-XC Leader□□□□□

Postgres-XL Leader□Mason Sharp

PGPool Leader□□□□□

PGStrom Leader□□□□□……

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□□PGer□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□PostgreSQL□

□□□□□□□□□□□□PostgreSQL 10□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□NoSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□PostgreSQL□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□PostgreSQL□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□

□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL
□□□□□□□□PostgreSQL□□□□□□□□□PGer□□□□

□□□□□□□□□□Postgres□□ □□

Postgres□□□□□2015-2018□ □□

□□□

2018□4□25□

[illegible]

PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]

[illegible]2010[illegible]PostgreSQL DBA[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL DBA[illegible]

[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]

[illegible]

[illegible]PostgreSQL[illegible]1[illegible]4[illegible]

PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□
SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□5□9
□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□10□18□□□□□□
PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
Oracle□□□□□PostgreSQL□□□PostGIS□□□□18
□□□□□□□□

□1□□□□PostgreSQL□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□

□2□□□□psql□□□□□□□□□□□□□□□□□psql□□
□□psql□□□□□□□□□psql□□□□□psql□□□□□□□

□3□□□□PostgreSQL□□□□□□□□□□□□□□□□□/
□□□□□□□□□□□□□□□□□□□□□□□□□□□□json/jsonb
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□4□□□□□□PostgreSQL□□□□□□□SQL□□□□
□WITH□□□□□□□□RETURNING□□DML□□□□□□
UPSERT□□□□□□□□□□□□□□□□□

□5□□□□□□PostgreSQL□□□□□□□□□□□□□
PostgreSQL□□□□□□□□□□□□□□□□

[illegible]6[illegible]PostgreSQL[illegible]

[illegible]7[illegible]PostgreSQL[illegible]

[illegible]8[illegible]

[illegible]9[illegible]PostgreSQL[illegible]NoSQL[illegible]PostgreSQL[illegible]

[illegible]10[illegible]Linux[illegible]

[illegible]11[illegible]PostgreSQL[illegible]pgbench[illegible]pgbench[illegible]

[illegible]12[illegible]PostgreSQL[illegible]

[illegible]13[illegible]PostgreSQL[illegible]

[illegible]14[illegible]Pgpool-II[illegible]Keepalived[illegible]

□□

□15□□□□PostgreSQL□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□16□□□□□□□□□□□□□□□□□□□file_fdw□
pg_stat_statements□auto_explain□
postgres_fdw□□□□□□Citus□□□□□

□17□□□□□□□□□□□□□□□□Oracle□□□□□□
PostgreSQL□□□□□□□□□

□18□□□□□□PostGIS□□□□□□□□□□□□□□□□□□
□□□□□□□□PostGIS□□□□□□□□□□□□□□□□□□□

□□□□

□□□□PostgreSQL□□□□□□□□□□PostgreSQL
□□□□□□□□□□□□PostgreSQL□□□□□□□□□□
PostgreSQL□□□□□□□□□□□□□□□□□□□□□□
PostgreSQL□□□□□□□□□□

□□□□PostgreSQL 10□□□□□□□□□□
PostgreSQL 10□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
PostgreSQL 10□□□□□□□

[illegible]18[illegible]PostgreSQL[illegible]
[illegible]
PostgreSQL[illegible]
PostgreSQL[illegible]

[illegible]

[illegible]PostgreSQL[illegible]

·PostgreSQL[illegible]DBA[illegible]PostgreSQL DBA[illegible]PostgreSQL[illegible]

·[illegible]PostgreSQL DBA[illegible]Oracle[illegible]MySQL[illegible]DBA[illegible]PostgreSQL[illegible]PostgreSQL[illegible]

·[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]

·[illegible]PostgreSQL[illegible]PostgreSQL[illegible]

[illegible]

[illegible]
francs.tan@postgres.cn[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]francs[illegible]

2018[illegible]4[illegible]

[illegible]

□1□ □□□□□□□

□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□

1.1 □□PostgreSQL

PostgreSQL□□PostgreSQL□□□□□□□□□□□□□□□□□□-□□□□□□□□□UC Berkeley□□1977□□Ingres□□□□□□□□□□□□□□□□Michael Stonebraker□2015□□□□□□□□□□□□1994□□□□UC Berkeley□□□□□□Andrew Yu□Jolly Chen□□□□□SQL□□□□□□□□□□□□□Ingres□QUEL□□□□□□Postgres95□□□□□□□□□□SQL□□□□□□□Postgres95□1996□□□□□PostgreSQL□□□□□□□□PostgreSQL□□□6.0□□□□2005□□PostgreSQL□□□□□□□□□□□□Windows□□□□8.0□□□□□2010□PostgreSQL 9.0□□□□PostgreSQL□□□□□□□□□□□□□□PostgreSQL□□□□□□□PostgreSQL 10□

PostgreSQL□□□PostgreSQL□RDS□□□

□□ PostgreSQL□□□□□□□□"Post-Gres-Q-L"□□□□`[/postgrs kju□l/]□`□□□□□□□□PostgreSQL□Postgres□□□□□□□□□□□□□□□□□□

□□□□□□□□□□PostgreSQL□□□□□□□□□□
http://www.postgresql.org/files/postgresql.mp3

1.1.1 PostgreSQL□□□

PostgreSQL□□□□□□□□□□□□□□□Linux□□□□□□UNIX□□UNIX□□□□Windows□□□□□AIX□BSD□HP-UX□SGI IRIX□Mac OS X□Solaris□Tru64□□□□□□□□□□□□C□C++□Go□Java□Perl□Python□Ruby□Tcl□□□□□□□□□ODBC□□□□□□□

□□□□□□□□□□□□□json□jsonb□□□□□□□□□□□SQL□□CREATE TYPE□□□□□□□□

□□□□□□SQL□□□□□□□□□SQL□□□□□SQL□□□□Window Function□□C□Java□python□R□□□□

□□□□□□□□□MVCC□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Hot Standby□□□□□□□□□□□□□Foreign data wrappers□□□□□□□□□□□□□□□□□□□□□□□Oracle□MySQL□Informix□SQLite□MS SQL Server□□

1.1.2 [illegible]

PostgreSQL[illegible]PostgreSQL License[illegible]BSD[illegible]MIT[illegible]OSI[illegible]PostgreSQL[illegible]PostgreSQL[illegible]

[illegible]https://www.postgresql.org/about/licence/[illegible]

1.1.3 [illegible]

PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]pgsql-zh-general[illegible]

PostgreSQL[illegible]https://www.postgresql.org/list [illegible]

1.2 □□PostgreSQL

PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□32□□64□□□□□□□□□□64□□CentOS 6□PostgreSQL 10□□□□□□□□□□□□□□□□□□

□□PostgreSQL□□□□□□□□□□yum□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

1.2.1 □□yum□□□

□□yum□□yum□□yum□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□yum□□□□□□□□□□

1.□□PostgreSQL□repository RPM

□□PostgreSQL□□□□ https://www.postgresql.org/download □□□□□□□□□□□□□□□□□□□□□□CentOS□□□□□□□□□“Binary packages”□□“Red Hat family Linux”□□□□□□□□□□Select version□□“10”□Select platform□□“CentOS 6”□Select

architecture□□“x86_64”□□□□□□□□□□□□□□□□
□□□□□□□□PostgreSQL□repository RPM□

```
[root@pghost1 ~]$ yum install
https://download.postgresql.org/pub/repos/yum/10/redhat/rhel
-6-x86_64/pgdg-centos10-10-1.noarch.rpm
```

□□□□□□□/etc/yum.repos.d□□□□□□□□□□
pgdg-10-centos.repo□□□□□□□

2.□□PostgreSQL

□□□PostgreSQL□repository RPM□□□□
yum□search□□□□□□□□□postgresql10□□□

```
[root@pghost1 ~]$ yum search postgresql10
```

□□□

·postgresql10-debuginfo.x86_64□
postgresql10□□□□□□□□□□□□DEBUG□□□□□□□
□□□□□□□□□□□□□□

·postgresql10.x86_64□□□□□□□
PostgreSQL□client□□□□□□□□□□□□□□□□□□□□

·postgresql10-contrib.x86_64□PostgreSQL□□□□□□□□□□□□□□□

·postgresql10-devel.x86_64□PostgerSQL□C□C++□□□□□□□□libpq□□□□□□□□□

·postgresql10-docs.x86_64□□□□

·postgresql10-server.x86_64□PostgreSQL server□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□server□contrib□□□□□□□□client□□□□□□□□□□□□□□□□□□□□□□

```
[root@pghost1 ~]$ yum install postgresql10-server
postgresql10-contrib
```

□□ □□□□□□□□□□□□□□yum□“-y”□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
[root@pghost1 ~]$ yum install -y postgresql10-server
postgresql10-contrib
```

□□□□□□□□□□□□□□□□□□□□□□□□□□yum□□□□□□□/usr/pgsql-10□□□□□□□□□□/usr/pgsql-10/bin□□□□□□□□□□□□postgres□□□□□home□□□/var/lib/pgsql□

3.□□□□yum□□□□PostgreSQL

□□□□□□□□□□□□□□□PostgreSQL□□□□

```
[root@pghost1 ~]$ rpm -qa | grep postgresql
postgresql10-10.0-1PGDG.rhel6.x86_64
postgresql10-libs-10.0-1PGDG.rhel6.x86_64
postgresql10-contrib-10.0-1PGDG.rhel6.x86_64
postgresql10-server-10.0-1PGDG.rhel6.x86_64
```

□□□□yum remove□□□□□□□□□□□□□□□□□□libs□□□□□□□□□□□□□□□□□□□libs□□□□□□□□□□□

```
yum remove postgresql10-libs-10.0-1PGDG.rhel6.x86_64
```

□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□

```
[root@pghost1 ~]$ rm -f /etc/init.d/postgresql-10
```

1.2.2　□□□□□□□□

□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□
□□

1.□□□□

□PostgreSQL□□□□
https://www.postgresql.org/ftp/latest □□□□
□□□□□□□□□□□□□

```
[root@pghost1 ~]$ wget
https://ftp.postgresql.org/pub/source/v10.0/postgresql-
10.0.tar.gz
```

□□□□□□□

```
[root@pghost1 ~]$ tar -xvf postgresql-10.0.tar.gz
```

2.□□configure□□□□□□□□

□□configure□□□□□□□□□□□□□□□□□□□□□
□□

```
[root@pghost1 ~]$ yum groupinstall "Development tools"
[root@pghost1 ~]$ yum install -y bison flex readline-devel
zlib-devel
```

□□□□□□□□□configure--help□□□□□□□□□□□□□□

```
[root@pghost1 ~]$ cd postgresql-10.0
[root@pghost1 ~]$ ./configure --help | less
```

PostgreSQL□□□□□□□□□□□□□□□□□□□

·--prefix=PREFIX□□□□□□□□□□□□□□□□“/usr/local/pgsql”□

·--includedir=DIR□□□C□C++□□□□□□□□□□□□□□□“PREFIX/include”□

·--with-pgport=PORTNUM□□RPM□□□□□□□□□□□□□□□

·--with-blocksize=BLOCKSIZE□□□□□□□□□□□□□□□□8kB□□□□OLAP□□□□□□□□□□□□□32kB□□□□OLAP□□□□□□OLTP□□□□□□□8kB□□□□

·--with-segsize=SEGSIZE□□□□□□□□□□□□□□□□1GB□

·--with-wal-blocksize=BLOCKSIZE□□□
WAL□□□□□□□□□□8kB□

·--with-wal-segsize=SEGSIZE□□□□□
WAL□□□□□□□□□16MB□

□□“--with-xxx-size”□4□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□configure□□□□□□□□□□□

```
[root@pghost1 postgresql-10.0]$ ./configure --
prefix=/opt/pg10/ --with-pgport=1921
checking build system type... x86_64-pc-linux-gnu
checking host system type... x86_64-pc-linux-gnu
...
checking for bison... /usr/bin/bison
configure: using bison (GNU Bison) 2.4.1
checking for flex... /usr/bin/flex
configure: using flex 2.5.35
...
configure: using CPPFLAGS= -D_GNU_SOURCE
configure: using LDFLAGS=  -Wl,--as-needed
configure: creating ./config.status
...
config.status: linking src/include/port/linux.h to
src/include/pg_config_os.h
config.status: linking src/makefiles/Makefile.linux to
src/Makefile.port
```

□□□configure□□□□□□□□□□□□
□“configure□error□readline library not

found”□□□□□□□□□□□□□□□□□□□□yum□□□□□□□

3.□□□□

□Linux□□PostgreSQL□□□□□□□□GNU make□□□□□□□gmake□□□□□□□gmake install□□gmake world□□□gmake install-world□□world□□

□□gmake□gmake world□□□□□□□□□□□□

```
[root@pghost1 ~]$ gmake
```

□□□□gmake□□□□□□□□□□□□□□□□□“All of PostgreSQL successfully made.Ready to install.”□□□□□□□□□

□□□□gmake world□□□□□□□□□□□□□□□□“PostgreSQL□contrib□and documentation successfully made.Ready to install.”□□□□□□□□□

□□gmake install□gmake install-world□□□□□□□□□□□□

```
[root@pghost1 ~]$ gmake install
```

□□□□gmake install□□□□□□□□□□□□□□□□“PostgreSQL installation complete.”□□□□□□□□□

□□□□gmake install-world□□□□□□□□□□□□□□□□“PostgreSQL□contrib□and documentation installation complete.”□□□□□□□□□

□□□□□PostgreSQL□□□□□□□□□□

```
[root@pghost1 ~]# /opt/pg10/bin/postgres --version
postgres (PostgreSQL) 10.0
```

1.2.3 □□□□□□□

□□□□□□□□□□□□□□□□□□shell□Python□□□□□□□□□□□□□□□□□□/opt/pg9.x□□□

[illegible]/opt/pgsql[illegible]

```
[root@pghost1 opt]$ ln -s /opt/pg10 /opt/pgsql
[root@pghost1 ~]$ ll /opt/
drwxr-xr-x  6 root    root    4096 Oct 11 14:32 pg96
drwxr-xr-x  6 root    root    4096 Oct 13 17:43 pg10
lrwxrwxrwx  1 root    root      10 Oct 13 11:25 pgsql ->
/opt/pg10/
```

[illegible]

1.3 [illegible]

[illegible]PostgreSQL[illegible]

```
[postgres@pghost1 ~]$ tree -L 1 /opt/pgsql/
/opt/pgsql/
├── bin
├── include
├── lib
└── share
4 directories, 0 files
```

share[illegible]PostgreSQL[illegible]man[illegible]include[illegible]PostgreSQL[illegible]C[illegible]C++[illegible]bin[illegible]PostgreSQL[illegible]PostgreSQL[illegible]C/S[illegible]

1.3.1 [illegible]

[illegible]

1.[illegible]SQL[illegible]

clusterdb

clusterdb[illegible]SQL CLUSTER[illegible]PostgreSQL[illegible]clusterdb[illegible]

[illegible]

```
[postgres@pghost1 ~]$ /opt/pgsql/bin/clusterdb -h pghost1 -p
1921 -d mydb
```

reindexdb

reindexdb[illegible]SQL REINDEX[illegible]reindexdb[illegible]

[illegible]

```
[postgres@pghost1 ~]$ /opt/pgsql/bin/reindexdb -e -h pghost1
-p 1921 -d mydb
```

vacuumdb

vacuumdb[illegible]PostgreSQL[illegible]VACUUM[illegible]VACUUM FREEZE[illegible]VACUUM FULL[illegible]VACUUM ANALYZE[illegible]SQL[illegible]

VACUUM[illegible]
PostgreSQL[illegible]

[illegible]

```
[postgres@pghost1 ~]$ /opt/pgsql/bin/vacuumdb -h pghost1 -p
1921 mydb
```

vacuumlo

vacuumlo[illegible]

[illegible]

```
[postgres@pghost1 ~]$ /opt/pgsql/bin/vacuumlo -h pghost1 -p
1921 mydb
```

createdb[illegible]dropdb

[illegible]SQL[illegible]CREATE DATABAS[illegible]DROP DATABASE[illegible]

[illegible]pghost1[illegible]1921[illegible]newdb[illegible]

```
[postgres@pghost1 ~]$ /opt/pgsql/bin/createdb -h pghost1 -p
1921 newdb "New database."
```

□□□□newdb□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ /opt/pgsql/bin/dropdb -h pghost1 -p
1921 newdb
```

createuser□dropuser

□□□□□SQL□□CREATE USER□DROP USER□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□newuser□□□□□□□newuser□□□pg_monitor□□□□□□□□□1□□□

```
[postgres@pghost1 ~]$ /opt/pgsql/bin/createuser -h pghost1 -
p 1921 -c 1 -g pg_monitor -D -R -S -P -e newuser
Enter password for new role:
Enter it again:
CREATE ROLE newuser PASSWORD
'md518b2c3ec6fb3de0e33f5612ed3998fa4' NOSUPERUSER NOCREATEDB
NOCREATEROLE INHERIT LOGIN CONNECTION LIMIT 1 IN ROLE
pg_monitor;
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□--interactive□□□□□□□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ /opt/pgsql/bin/createuser -h pghost1 -
p 1921 -c 1 -g pg_monitor --interactive -e -P newuser
```

```
Enter password for new role:
Enter it again:
Shall the new role be a superuser? (y/n) n
Shall the new role be allowed to create databases? (y/n) n
Shall the new role be allowed to create more new roles?
(y/n) n
CREATE ROLE newuser PASSWORD
'md545c93e6e78f597d46a41cfb08dea5ae3' NOSUPERUSER NOCREATEDB
NOCREATEROLE INHERIT LOGIN CONNECTION LIMIT 1 IN ROLE
pg_monitor;
```

[illegible]newuser[illegible]

```
[postgres@pghost1 ~]$ /opt/pgsql/bin/dropuser -h pghost1 -p
1921 newuser
```

2.[illegible]

pg_basebackup[illegible]PostgreSQL[illegible]

pg_dump[illegible]pg_dumpall[illegible]

pg_restore[illegible]pg_dump[illegible]

[illegible]13[illegible]

3.[illegible]

ecpg□□□C□□□PostgreSQL□□□SQL□□□□□□□SQL□□□□□□□□□□□□□□□□□□SQL□□□C□□□□□□□C□□□□□□□□□□□□C□□□□□□□□

·oid2name□□□□PostgreSQL□□□□□□OID□□□□□□□□□□□□□□□□□□□□

·pgbench□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

·pg_config□□□□□□□PostgreSQL□□□□□□□□□□

·PostgreSQL□□□pg_isready□□□□□□□□□□□□□□□□□□□□□□□

·pg_receivexlog□□□□□□□□□□□□□□□□□□□□

·pg_recvlogical□□□□□□□□□□□□□□□□□□□□□□□

·psql□□□PostgreSQL□□psql□□□□□□□□□□□□□□□□□□

```
[postgres@pghost2 ~]$ /opt/pgsql/bin/psql -h pghost1 -p 1921
mydb
psql (10.0)
Type "help" for help.
mydb=#
```

[illegible]

·-h[illegible]

·-p[illegible]

·-d[illegible]

[illegible]PostgreSQL[illegible]shell[illegible]“=#”[illegible]“=>”[illegible]

[illegible]“\q”[illegible]CTRL+D[illegible]

```
[postgres@pghost2 ~]$ /opt/pgsql/bin/psql -h pghost1 -p 1921
mydb
psql (10.0)
Type "help" for help.
mydb=# \q
[postgres@pghost2 ~]$
```

psql[illegible]
psql[illegible]

1.3.2 [illegible]

[illegible]

·initdb[illegible]

·pg_archivecleanup[illegible]PostgreSQL WAL[illegible]

·pg_controldata[illegible]

·pg_ctl[illegible]

·pg_resetwal[illegible]pg_control[illegible]pg_resetwal[illegible]

·pg_rewind[illegible]master[illegible]slave[illegible]master[illegible]

·pg_test_fsync[illegible]wal_sync_method[illegible]I/O[illegible]

·pg_test_timing[illegible]

·pg_upgrade□PostgreSQL□□□□□□□□□□□□□□□□□□

·pg_waldump□□□□□□□□□□□□□□□

·postgres□PostgreSQL□□□□□□

·postmaster□□□bin□□□□□□□□postgres□□□□□□□□□□□

1.4 [illegible]

[illegible]PostgreSQL[illegible]PostgreSQL[illegible]

1.4.1 [illegible]

[illegible]BUG[illegible]pger[illegible]postgres[illegible]start[illegible]stop[illegible]restart[illegible]yum[illegible]postgres[illegible]postgres[illegible]postgres[illegible]yum[illegible]postgres[illegible]

[illegible]BUG[illegible]PostgreSQL[illegible]BUG[illegible]PostgreSQL[illegible]

□□□□□□□□□□□□□□PostgreSQL□□□BUG□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□

```
[root@pghost1 ~]$ groupadd -g 1000 postgres
[root@pghost1 ~]$ useradd -g 1000 -u 1000 postgres
[root@pghost1 ~]$ id postgres
uid=1000(postgres) gid=1000(postgres) groups=1000(postgres)
```

□□□□□

1□□□□□□□□□□□□□□□□□□□root□□□□□□□□□□□□□□□□□□□sudo□□□□□□

2□□□□□□□□□□□□□NTP□□□□□□□□□□□□□□□□□□□□uid□gid□□□□□□□□□□□□□□□□□□□□□uid□gid□□□□□□□□□□□□□groupmod□□□usermod□□□□□□□□□□

```
[root@pghost1 ~]$ groupmod -g 1000 postgres
[root@pghost1 ~]$ usermod -u 1000 -g 1000 postgres
```

1.4.2 □□□□□□

[illegible]SQL[illegible]

[illegible]/pgdata/9.x/xxx_data[illegible]9.x[illegible]10[illegible]xxx_data[illegible]xxx[illegible]/pgdata/10/data[illegible]data[illegible]backups[illegible]scripts[illegible]archive_wals[illegible]

```
[root@pghost1 ~]$ mkdir -p
/pgdata/10/{data,backups,scripts,archive_wals}
```

[illegible]0700[illegible]initdb[illegible]PostgreSQL[illegible]

```
[root@pghost1 ~]$ chown -R postgres.postgres /pgdata/10
[root@pghost1 ~]$ chmod 0700 /pgdata/10/data
```

1.4.3 □□□□□□□

□□□□□□□□□initdb□□□initdb□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□template1□postgres□□□□□□□□□□□□□□□□□□□□□□□□□initdb□□□□□□□□□□

```
      [postgres@pghost1 ~]$ /opt/pgsql/bin/initdb --help
      initdb initializes a PostgreSQL database cluster.
      Usage:
          initdb [OPTION]... [DATADIR]
      Options:
          -A, --auth=METHOD       □□□□□□□pg_hba.conf□□□□□□□□□□
  □□md5□
                                 trust□password□□□□□□□□□□□□□□
  trust□□
                                 □□□□□□□□□□□□□□□□□□□□□□□□□
              --auth-host=METHOD  □□□□TCP/IP□□□□□□□□
  pg_hba.conf□□□□□□                  □□□
              --auth-local=METHOD □□□□UNIX Socket□□□□□□□□
  pg_hba.conf□□□□
                                 □□□□□
         [-D, --pgdata=]DATADIR   □□□□□□□□□□□□□□□□□□□□□□□□□□□□
  □□□□
          -E, --encoding=ENCODING □□□□□□□□□□□□□□□□□□template1□
  □□□□□
                                 □□□□□□□□□□□□template1□□□□□□□
              --locale=LOCALE     □□□□
              --lc-collate=, --lc-ctype=, --lc-messages=LOCALE
              --lc-monetary=, --lc-numeric=, --lc-time=LOCALE
                                  □□□□□□□□□□
              --no-locale         □□□ --locale=C
              --pwfile=FILE       □□□□□□□□□□□□□□□□□□□□□□□
          -T, --text-search-config=CFG
                                  □□□□□□□□□□□□
          -U, --username=NAME     □□□□□□□□□□□□□□□□□postgres□
          -W, --pwprompt          □initdb□□□□□□□□□□□□□□□□□□□
          -X, --waldir=WALDIR     □□□□□□□WAL□□□□□□□
```

[illegible]

```
[postgres@pghost1 ~]$ /opt/pgsql/bin/initdb -D
/pgdata/10/data -W
The files belonging to this database system will be owned by
user "postgres".
This user must also own the server process.
The database cluster will be initialized with locale
"en_US.UTF-8".
The default database encoding has accordingly been set to
"UTF8".
The default text search configuration will be set to
"english".
Data page checksums are disabled.
Enter new superuser password:
Enter it again:
fixing permissions on existing directory /export/pg10_data
... ok
creating subdirectories ... ok
selecting default max_connections ... 100
selecting default shared_buffers ... 128MB
selecting dynamic shared memory implementation ... posix
creating configuration files ... ok
running bootstrap script ... ok
performing post-bootstrap initialization ... ok
syncing data to disk ... ok
WARNING: enabling "trust" authentication for local
connections
You can change this by editing pg_hba.conf or using the
option -A, or
--auth-local and --auth-host, the next time you run initdb.
Success. You can now start the database server using:
    /opt/pgsql/bin/pg_ctl -D /pgdata/10/data -l logfile
start
[postgres@pghost1 ~]$
```

[illegible]-W[illegible]initdb[illegible]initdb[illegible]

template1[illegible]postgres[illegible]template1[illegible]postgres[illegible]

[illegible]

```
initdb: directory "/pgdata/10/data" exists but is not empty
If you want to create a new database system, either remove
or empty
the directory "/pgdata/10/data" or run initdb
with an argument other than "/pgdata/10/data".
```

[illegible]

[illegible]initdb[illegible]pg_ctl[illegible]

```
[postgres@pghost1 ~]$ /opt/pgsql/bin/pg_ctl init -D
/pgdata/10/data -o "-W"
```

[illegible]

[illegible]yum[illegible]PostgreSQL[illegible]/var/lib/pgsql/10[illegible]data[illegible]backups[illegible]service postgresql-10init[illegible]

□□□/var/lib/pgsql/10/data□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□

1.5 [illegible]

[illegible]service[illegible]PostgreSQL[illegible]

1.5.1 [illegible]service[illegible]

[illegible]

```
[root@pghost1 ~]$ service postgresql-10 start
```

[illegible]

```
[root@pghost1 ~]$ service postgresql-10 status
```

[illegible]

```
[root@pghost1 ~]$ service postgresql-10 stop
```

1.5.2 [illegible]pg_ctl[illegible]

pg_ctl□PostgreSQL□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□service□
systemctl□□□□□□pg_ctl□□□□□□□□□□□□□
pg_ctl□□□□□□□□□□□□su□□□□□postgres□□□

1.□□□□□

□□□□□□□

```
[root@pghost1 ~]# su - postgres
[postgres@pghost1 ~]$ /opt/pgsql/bin/pg_ctl -D
/pgdata/10/data start
server started
```

2.□□□□□□□□□

□□□□□□□

```
[root@pghost1 ~]# su - postgres
[postgres@pghost1 ~]$ /opt/pgsql/bin/pg_ctl -D
/pgdata/10/data status
pg_ctl: no server running
```

□□□

```
pg_ctl: server is running (PID: 43965)
/opt/pgsql/bin/postgres "-D" "/pgdata/10/data"
```

□□□□□pg_isready□□□□□□□□□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ /opt/pgsql/bin/pg_isready -p 1921
/tmp:1921 - accepting connections
```

□□□

```
/tmp:1921 - no response
```

3.□□□□□

□□pg_ctl□□□□□□□□□□

```
pg_ctl stop     [-D DATADIR] [-m SHUTDOWN-MODE] [-W] [-t SECS] [-s]
```

“-s”□□□□□□□□□□□□□□□□“-t SECS”□□□□□□□□□□□SECS□□□□□□□□□□□□□□□□“-m”□□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□smart□fast□immediate□□□□fast□□□

·smart□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

·fast□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□

·immediate□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□“-ms”“-mf”“-mi”□□□□□smart□□□□□□□□□□□□□□□

```
[root@pghost1 ~]# su - postgres
[postgres@pghost1 ~]$ /opt/pgsql/bin/pg_ctl -D
/pgdata/10/data -ms stop
```

1.5.3 □□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□postmaster□postgres□□□□□□□□□□□□□□□

```
[root@pghost1 ~]# su - postgres
[postgres@pghost1 ~]$ /opt/pgsql/bin/postgres -D
/pgdata/10/data/
```

□□□□□□□□□□□□□□□□□□□“&”□□□□□□□□□□

□PostgreSQL□□□□□postmaster□□□□□□□□□□□□□□□□□SIGINT□SIGTERM□SIGQUIT□□□□

□□□□□PostgreSQL□□□□□□□smart□fast□immediate□□□□□□□□□□kill□□□postgres□□□□SIGTERM□SIGINT□SIGQUIT□□□□□□□□□□□□smart□□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ kill -sigterm `head -1
/pgdata/10/data/postmaster.pid`
received smart shutdown request
shutting down
database system is shut down
```

□□□□□□□□□□□□□□□□smart□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□pqsignal□pmdie□□□□□□□□□□

□□PostgreSQL□□□□□□□□□□□pg_ctl□□□□□□□kill□□□□□□□□□□□□□

1.5.4 □□□□□□

□□□□□□yum□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

1.□□□□□□

□□□□□contrib□□□□Linux□FreeBSD□OSX□□□□□□□□□□□□□

```
[root@pghost1 ~]$ ls postgresql-10.0/contrib/start-scripts/
freebsd  linux  osx
```

[illegible]linux[illegible]/etc/init.d/[illegible]postgresql-10[illegible]

```
[root@pghost1 ~]$ cp postgresql-10.0/contrib/start-
scripts/linux /etc/init.d/postgresql-10
[root@pghost1 ~]$ chmod +x /etc/init.d/postgresql-10
[root@pghost1 ~]$ ls -lh /etc/init.d/postgresql-10
-rwxr-xr-x 1 root root 3.5K Oct 13 16:30
/etc/init.d/postgresql-10
```

2.[illegible]

chkconfig--list[illegible]PostgreSQL[illegible]

```
[root@pghost1 ~]$ chkconfig --list | grep postgresql-10
postgresql-10   0:off   1:off   2:off   3:off   4:off
5:off   6:off
```

chkconfig[illegible]PostgreSQL[illegible]

```
[root@pghost1 ~]$ chkconfig postgresql-10 on/off
```

1.6 [illegible]

[illegible]Database[illegible]

PostgreSQL[illegible]postgresql.conf[illegible]pg_hba.conf[illegible]postgresql.conf[illegible]pg_hba.conf[illegible]

1.6.1 [illegible]

[illegible]postgresql.conf[illegible]postgresql.auto.conf[illegible]pg_hba.conf[illegible]pg_ident.conf[illegible]postgresql.conf[illegible]

1.6.2 pg_hba.conf

pg_hba.conf[illegible]“[illegible]”[illegible]

```
TYPE  DATABASE        USER            ADDRESS
METHOD
local database user auth-method [auth-options]
host database user address auth-method [auth-options]
hostssl database user address auth-method [auth-options]
hostnossl database user address auth-method [auth-options]
host database user IP-address IP-mask auth-method [auth-
options]
hostssl database user IP-address IP-mask auth-method [auth-
options]
hostnossl database user IP-address IP-mask auth-method
[auth-options]
```

[illegible]ADDRESS[illegible]TYPE[illegible]METHOD[illegible]USER[illegible]DATABASE[illegible]

1.[illegible]

TYPE[illegible]local[illegible]host[illegible]hostssl[illegible]hostnossl[illegible]

·local[illegible]Unix[illegible]TYPE[illegible]local[illegible]Unix[illegible]

·host[illegible]TCP/IP[illegible]SSL[illegible]SSL[illegible]localhost[illegible]TCP/IP[illegible]postgresql.conf[illegible]listen_addresses[illegible]

·hostssl□□□□□□□SSL□TCP/IP□□□□□hostssl□□□□□□□□

1.□□□□□□□□□□OpenSSL□

2.□□PostgreSQL□□□□□configure□□--with-openssl□□SSL□□□

3.□postgresql.conf□□□ssl=on□

·hostnossl□hostssl□□□□□□□□□□SSL□TCP/IP□□□

2.□□□□□

DATABASE□□□□□□□□□□□□□□□□

3.□□□□

USER□□□□□□□□□□□□□□□□□□

4.□□□□

ADDRESS□□□□□□□□□□IP□□□IP□□□□□□

5.□□□□

METHOD[illegible]trust[illegible]reject[illegible]md5[illegible]password[illegible]

reject[illegible]

md5[illegible]password[illegible]md5[illegible]md5[illegible]password[illegible]password[illegible]

scram-sha-256[illegible]PostgreSQL 10[illegible]SASL[illegible]PostgreSQL[illegible]scram-sha-256[illegible]PostgreSQL 10[illegible]

```
[postgres@pghost2 ~]$ /usr/pgsql-9.6/bin/psql -h pghost1 -p
1921 -U postgres mydb
psql: SCRAM authentication requires libpq version 10 or
above
```

[illegible]https://www.postgresql.org/docs/current/static/auth-methods.html [illegible]

1.6.3 postgresql.conf

postgresql.conf□□□□□□□□□□□□□□□□configparameter=value□□□□□□□“#”□□□□□□□□value□□□□□□□□□□□□□□□□□□□□□□□□□value□□□□□□□□□□□□MB□GB□ms□min□d□□postgresql.conf□□□□□include□include_if_exists□□□□□□□□□□

□□□□□□□□□“#□change requires restart□”□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□reload□□□□□

1.□□□□□□□□□

□□□□□□□□□□□

·□□postgresql.conf□□□□□

·□□vim□nano□□□□□□□□□sed□□□□□□□

·□□ALTER SYSTEM□□□□□□□□□□□□

```
mydb=# ALTER SYSTEM SET listen_addresses = '*';
```

□□ALTER SYSTEM SQL□□□□□□□□□□□□□□□□□postgresql.auto.conf□□□□□□□□□□□□□

postgresql.auto.conf□□□□□□□□□□□
postgresql.conf□□□□□□□□□□□□□□□□□□□

·□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ /opt/pgsql/bin/postgres -D
/pgdata/10/data -c port=1922
```

2.□□□□□□□□□□

·□□□□□Database□□□□□□□□□

```
ALTER DATABASE name SET configparameter { TO | = } { value |
DEFAULT }
ALTER DATABASE name RESET configuration
```

·□□□□□Session□□□□□□

·□□SET□□□□□□Session□□□□□□□

```
SET configparameter { TO | = } { value | 'value' | DEFAULT }
SET configparameter TO DEFAULT;
```

·□□pg_settings□□□□□□

```
UPDATE `pg_settings` SET setting = new_value WHERE name =
'configparameter';
```

```
UPDATE `pg_settings` SET setting = reset_val WHERE name =
'configparameter';
```

·□□set_config□□□□□□□□□□□□

```
SELECT set_config('configparameter',new_value,false);
```

·□□□□□Role□□□□□□□□□

```
ALTER ROLE name IN DATABASE database_name SET
configparameter { TO |= } { value | DEFAULT }
    ALTER ROLE name IN DATABASE database_name RESET
configparameter
```

3.□□□□□□

□□pg_settings□□□□□□□

```
SELECT name,setting FROM pg_settings where name ~ ‘xxx’;
SELECT current_setting(name);
```

□□show□show all□□□□□□

4.□□□□□□□□

□□□□□□□□□□□□reload□□□□□□□□□□□□□□□

```
mydb=# SELECT pg_reload_conf();
pg_reload_conf
----------------
t
(1 row)
```

□□□□□pg_ctl□□reload□□□□□□□□□□

```
[root@pghost1 ~]# su - postgres
[postgres@pghost1 ~]$ /opt/pgsql/bin/pg_ctl -D
/pgdata/10/data reload
```

1.6.4 □□□□□□□□

□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ netstat -nlt | grep 1921
Active Internet connections (only servers)
Proto Recv-Q Send-Q Local Address           Foreign Address
State
tcp        0      0 127.0.0.1:1921            0.0.0.0:*
LISTEN
tcp        0      0 ::1:1921                  :::*
LISTEN
```

□□□□□□□□□□□□□□□□□□□□□□□□

```
[postgres@pghost2 ~]$ telnet pghost1 1921
Trying pghost1...
telnet: connect to address pghost1: Connection refused
```

□□□□□□□□□□□□□□□□□□□□

1.□□□□□□

PostgreSQL□□□□□□□□□□□postgresql.conf□□□□listen_addresses□□□□□□□□□□□□□□localhost□□□□□□□□□TCP/IP□□□□□□□□□□□□□□□□□postgresql.conf□□listen_addresses□□□□□□□□□□□postgresql.conf□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ vim /pgdata/10/data/postgresql.conf
```

□□□□□listen_addresses□□□□□□□□□□

```
#listen_addresses = 'localhost'   # what IP address(es) to
listen on;
    # comma-separated list of addresses;
    # defaults to 'localhost'; use '*' for all
    # (change requires restart)
```

□□listen_addresses□□□4□□□□□□□□□□

·what IP address□es□to listen on—□□□□IP□□□□□□□□□□□IP□□□□□□□□□□IP□□□□□□□IP□□□

·comma-separated list of addresses□—□□□□□□□□□□

·defaults to'localhost'□use'*'for all—□□□“localhost”□□□“*”□□□□□□□□□□□□□□□□□□□VIP□□□□□□□□□□□□□□□□□□“*”□

·□change requires restart□—□□□□□□□□□□□□□□□□

□□listen_addresses□□□□□□“#”□□□□□□□□□□“*”□□□□□□□□□□□□□□□□□□□

```
listen_addresses = '*'
```

□□□□□□□□□□□□□□□□□□□□□□

```
[root@pghost1 ~]$ service postgresql-10 restart
```

□□□□□□□□□□□□□□pg_hba.conf□□□□□pg_hba.conf□□TCP/IP□□□□□□□□□□SSL□□SSL□□□□□md5□□□□□□□□pguser□□□□□mydb□□□□□□□□□□□□pg_hba.conf□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ echo "host mydb pguser 0.0.0.0/0 md5"
>> /pgdata/10/data/pg_hba.conf
```

□□pg_hba.conf□□□□□□reload□□□□□□□□□□

```
[postgres@pghost1 ~]$ /opt/pgsql/bin/pg_ctl -D
/pgdata/10/data/ reload
server signaled
2017-10-18 10:16:00.405 CST [36171] LOG:  received SIGHUP,
reloading configuration files
```

□□□□□□□□□□□□□□□□

□□Windows□□□□Linux□□□selinux□iptables□□□□□□□□□□Linux□□□□□□□selinux□□□iptables□□□□□□□□□□□□□□□□iptables□□□□□□□□□□□□□□□□□□□□□□

1.7 [illegible]

[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]

□2□ □□□□□

□□□□□PostgreSQL□□□□□□□□pgAdmin□psql□pgAdmin□□□□□□□□□□□□□PostgreSQL□□□□□□□□□psql□PostgreSQL□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

2.1 pgAdmin 4□□

pgAdmin□□□□□PostgreSQL□□□□□□□□□□□□□□□https://www.pgadmin.org/ □□□pgAdmin 4□□□□□□□□□□□□□□□□□

2.1.1 pgAdmin 4□□

pgAdmin□□Linux□Unix□Mac OS X□Windows□□□□□□□□pgAdmin□□□□□□□4□□□□□pgAdmin□□□□□□pgAdmin 4□□□□□Windows 7□□□pgAdmin 4□□□□□□□pgAdmin 4□

□□□□□□□□□https://www.postgresql.org/ftp/pgadmin/pgadmin4/v1.6/windows /□□□□□□□□□□□□□□□□□□□pgAdmin 4□□□□□2-1□□□

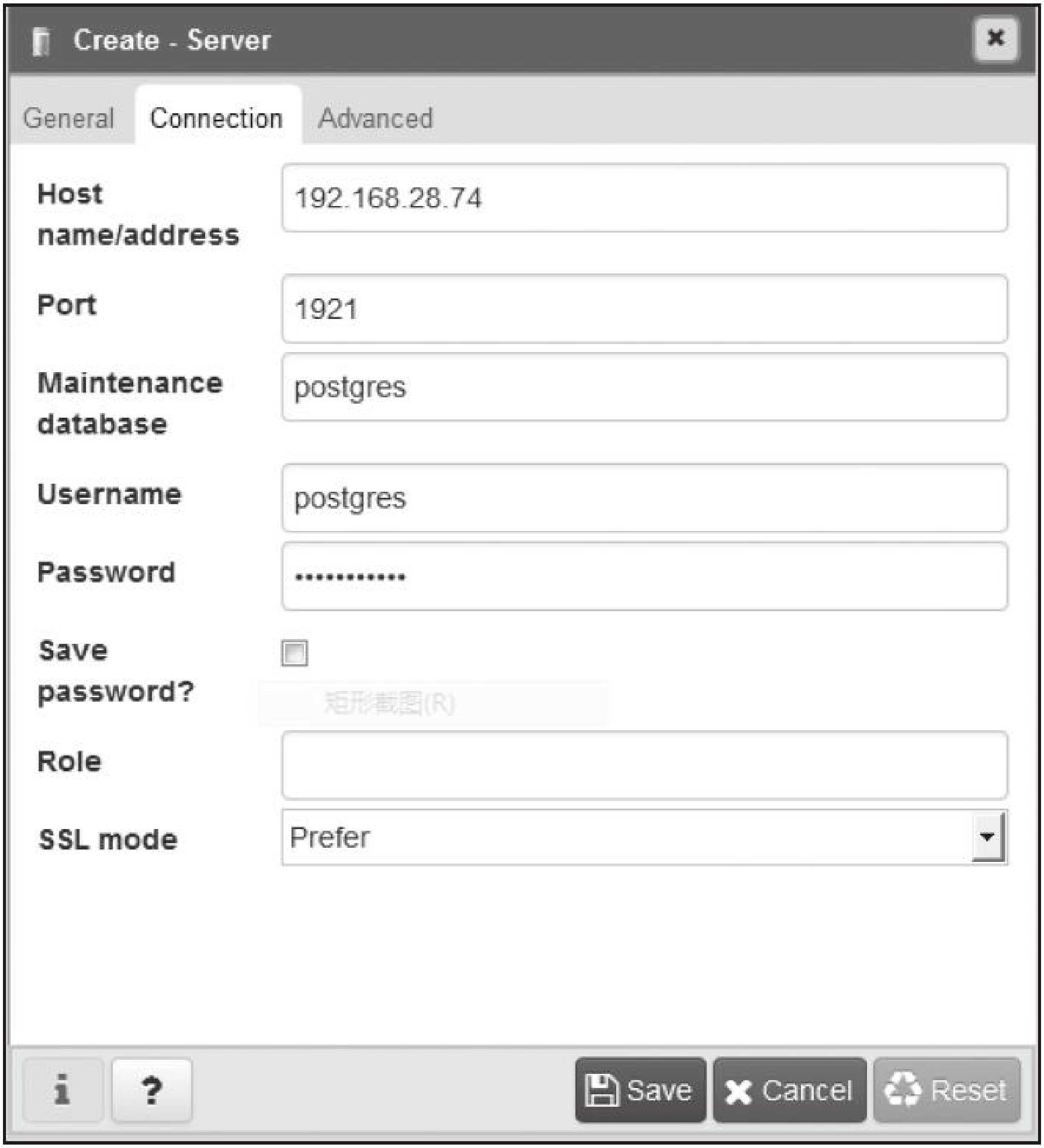
2.1.2 pgAdmin 4□□

pgAdmin 4□□□□□□□□□□□□□□□□□□□pgAdmin 4□□PostgreSQL□□□□□□□□□□□□□

□2-1　pgAdmin 4□□

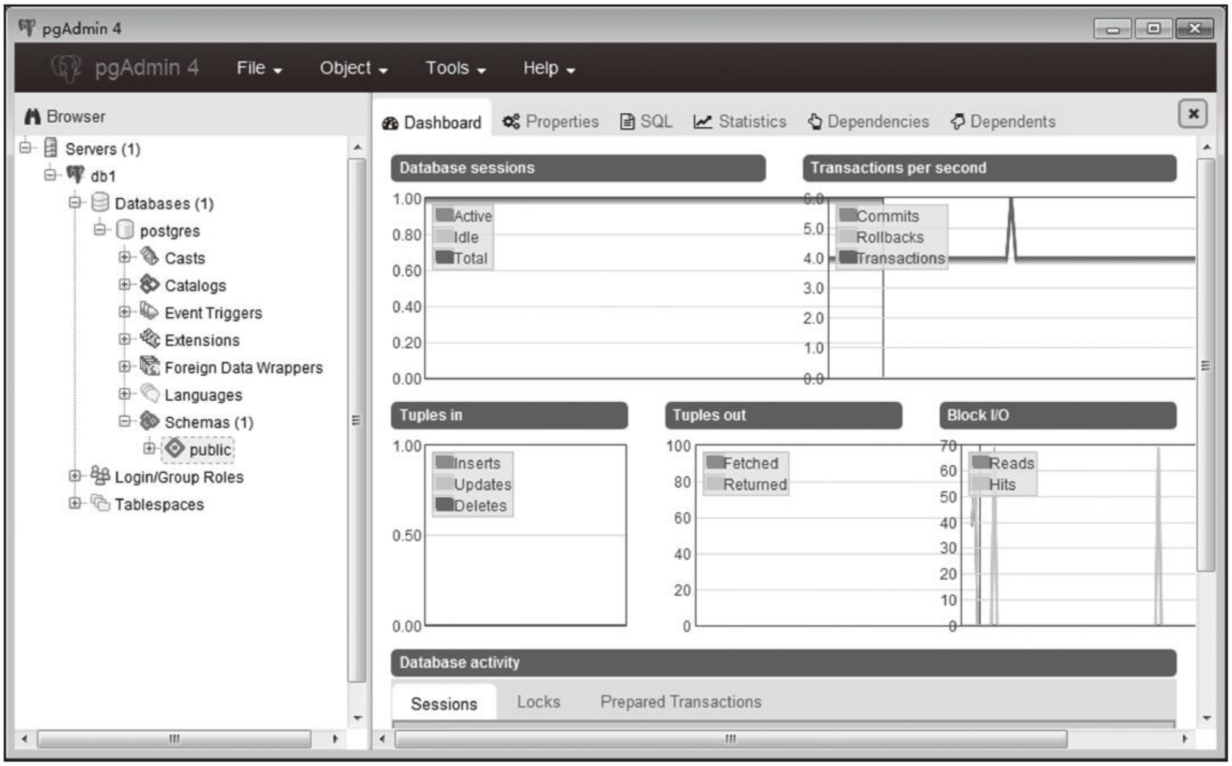
1.pgAdmin 4□□□□□

□□pgAdmin 4□□□□□□□□□□□□□□□□□□□□□2-2□□□



□2-2 □□pgAdmin 4□□□□□

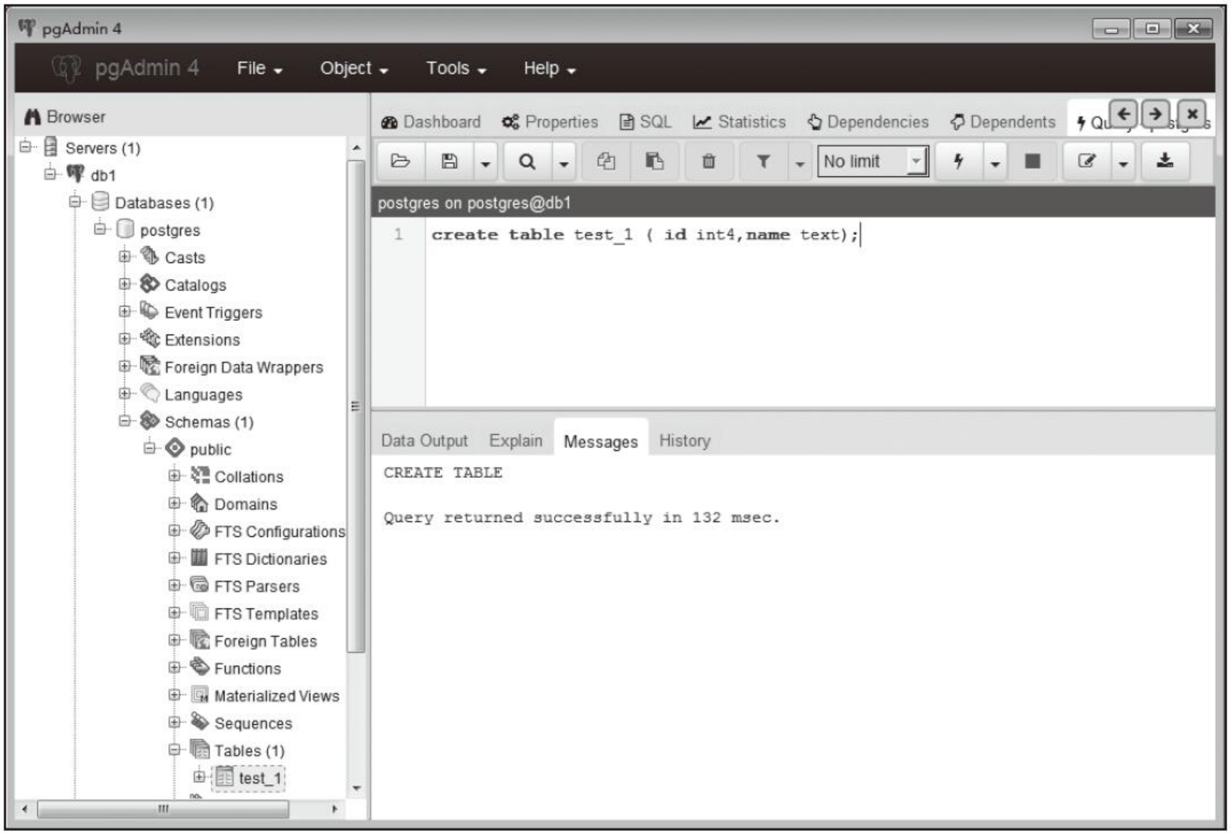
□General□□□□□□□□□□□□□□□□□□□db1□Connection□□□□□□□□□□□□□□2-3□□□



□2-3　□□pgAdmin 4□□□□□

2.pgAdmin 4□□□□□□□

□pgAdmin 4□□□□□Tools□□□□Query Tool□□□□□□□□□□□□□□□□□□DDL□DML□□□□□□□□□□□□□□□2-4□□□

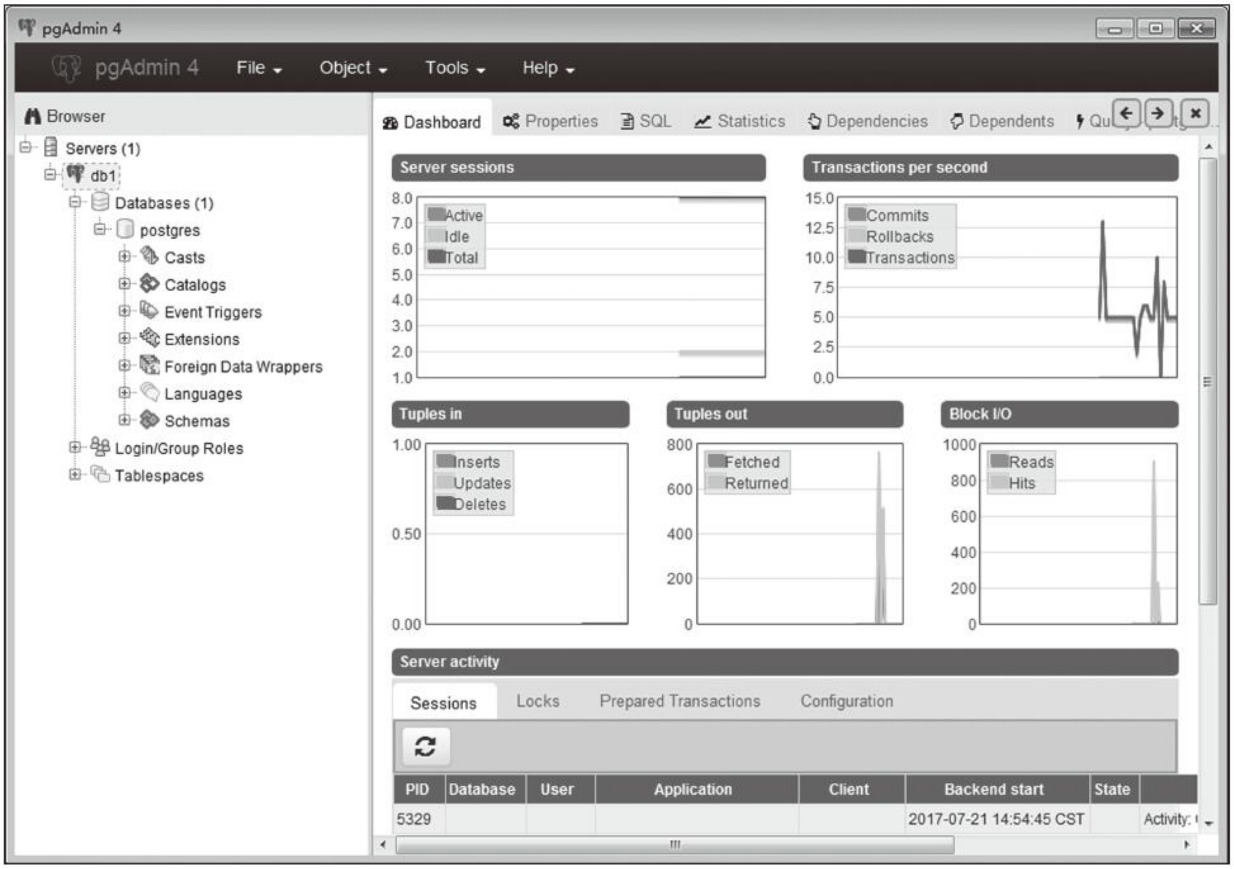


□2-4　□□pgAdmin 4□□□

□□□□□□□pgAdmin 4□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□DDL□□□□□□□□□□

3.□pgAdmin 4□□□□□□

pgAdmin 4□□□□□□□□□□□□2-5□□□□□□□□□□□□□□□□□□□□□□□□□□□□□



□2-5　□□pgAdmin 4□□□□□□□□

pgAdmin 4□□□□□□□□pgAdmin 4□□□□□
□□□□□□□□□

2.2 psql□□□□□

psql□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□Oracle□□□□□□□sqlplus□□□□□□psql□□□□□□□□□□□□□□□psql□□□□PostgreSQL□□□□□□

2.2.1 □□psql□□□□

□psql□□

```
[postgres@pghost1 ~]$ psql postgres postgres
psql (10.0)
Type "help" for help.
postgres=#
```

psql□□□□□postgres□□□□□□□postgres□□□□□□□□□□□□□$PGPORT□□□□□□□□□□□1921□□□□□□□□□□□□□□□□mydb□□□□□pguser□□□mydb□□□□□□□□tbs_mydb□□□□□

```
--□□□□
postgres=# CREATE ROLE pguser WITH ENCRYPTED PASSWORD
'pguser';
```

```
CREATE ROLE

--□□□□□□□
[postgres@pghost1 ~]$ mkdir -p
/database/pg10/pg_tbs/tbs_mydb

--□□□□□
postgres=# CREATE TABLESPACE tbs_mydb OWNER pguser LOCATION
'/database/pg10/pg_tbs/tbs_mydb';
CREATE TABLESPACE

--□□□□□
postgres=# CREATE DATABASE mydb
             WITH  OWNER = pguser
             TEMPLATE = template0
             ENCODING = 'UTF8'
             TABLESPACE = tbs_mydb;
CREATE DATABASE

--□□
GRANT ALL ON DATABASE  mydb TO pguser WITH GRANT OPTION;
GRANT ALL ON TABLESPACE tbs_mydb TO pguser;
```

CREATE DATABASE□□□□OWNER□□□□□□□□□□TEMPLATE□□□□□□□□□□□template0□template1□□□□□□□□□□□□□□ENCODING□□□□□□□□□□□□□□UTF8□TABLESPACE□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□psql□□□□□□□□□□□□□□□

```
psql [option...] [dbname [username]]
```

□□□pghost1□IP□192.168.28.74□pghost2□IP□192.168.28.75□□pghost2□□□□

□□□pghost1□□mydb□□□□□□

```
[postgres@pghost2 ~]$ psql -h 192.168.28.74 -p 1921 mydb
pguser
Password for user pguser:
psql (10.0)
Type "help" for help.
```

□□psql□□□□□□□\q□□□□CTRL+D□□□□□□□
□□□□

```
[postgres@pghost1 ~]$ psql mydb pguser
psql (10.0)
Type "help" for help.
mydb=> \q
```

□□□□□□□□□psql□□□□□□□

2.2.2 psql□□□□□

psql□□□□□□□□□□□□□□□□□psql□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□

```
postgres=# \l
                                 List of databases
   Name    |  Owner   | Encoding | Collate | Ctype |
Access privileges
```

```
-----------+----------+----------+---------+-------+--------
---------------
 mydb      | postgres | UTF8     | C       | C     |
=Tc/postgres         +
           |          |          |         |       |
postgres=CTc/postgres+
           |          |          |         |       |
pguser=C/postgres
 postgres  | postgres | UTF8     | C       | C     |
 template0 | postgres | UTF8     | C       | C     |
=c/postgres          +
           |          |          |         |       |
postgres=CTc/postgres
 template1 | postgres | UTF8     | C       | C     |
=c/postgres          +
           |          |          |         |       |
postgres=CTc/postgres
(4 rows)
```

1..\db[illegible]

[illegible]\db[illegible]

```
postgres=# \db
                  List of tablespaces
    Name    |  Owner   |             Location
------------+----------+----------------------------------
pg_default | postgres |
pg_global  | postgres |
tbs_mydb | pguser   | /database/pg10/pg_tbs/tbs_mydb
(3 rows)[illegible]
```

2.\d[illegible]

[illegible]

```
mydb=> CREATE TABLE test_1(id int4,name text,
    create_time timestamp without time zone default
clock_timestamp());
CREATE TABLE

mydb=> ALTER TABLE test_1 ADD PRIMARY KEY (id);
ALTER TABLE
```

generate_series[illegible]test_1[illegible]\d[illegible]

```
mydb=> \d test_1
                              Table "pguser.test_1"
    Column |            Type            | Collation |
Nullable |      Default
-----------+----------------------------+-----------+-------
---+-----------------
id         | integer                    |           | not
null |
name       | text                       |           |
|
create_time| timestamp without time zone |           |
| clock_timestamp()
Indexes:
    "test_1_pkey" PRIMARY KEY , btree (id)
```

3.[illegible]

[illegible]test_1[illegible]500[illegible]

```
mydb=> INSERT INTO test_1(id,name)
       SELECT n,n || '_francs'
       FROM generate_series(1,5000000) n;
INSERT 0 5000000
```

□□□□□□□\dt+□□□□□□□□□□

```
mydb=> \dt+ test_1
                   List of relations
Schema |  Name  | Type  | Owner  |  Size  | Description
-------+--------+-------+--------+--------+-------------
pguser | test_1 | table | pguser | 287 MB |
(1 row)
```

□□□□□□□□\di+□□□□□□□□□□□

```
mydb=> \di+ test_1_pkey
                          List of relations
Schema |    Name     | Type  | Owner  | Table  |  Size  |
Description
-------+-------------+-------+--------+--------+--------+---
----------
pguser | test_1_pkey | index | pguser | test_1 | 107 MB |
(1 row)
```

4.\sf□□□□□□

□□□\sf□□□□□□□□□□□□□□□□□□

```
mydb=> \sf random_range
CREATE OR REPLACE FUNCTION pguser.random_range(integer,
integer)
    RETURNS integer
    LANGUAGE sql
AS $function$
    SELECT ($1 + FLOOR(($2 - $1 + 1) * random() ))::int4;
    $function$
```

[illegible]\sf[illegible][illegible]random_range[illegible]integer[illegible]integer[illegible]PostgreSQL[illegible][illegible]\sf[illegible]

5.\x[illegible]

[illegible]\x[illegible]

```
mydb=> SELECT * FROM test_1 LIMIT 1;
    id |   name   |        create_time
-------+----------+---------------------------
     1 | 1_pguser | 2017-07-22 11:16:15.97559
(1 row)

mydb=> \x
Expanded display is on.
mydb=> SELECT * FROM test_1 LIMIT 1;
-[ RECORD 1 ]--------------------------
id          | 1
name        | 1_francs
create_time | 2017-07-22 11:16:15.97559
```

6.[illegible]SQL[illegible]

psql[illegible]SQL[illegible]psql[illegible]-E[illegible]SQL[illegible]

```
[postgres@pghost1 ~]$ psql -E mydb  pguser
psql (10.0)
Type "help" for help.
```

```
mydb=> \db
********* QUERY **********
SELECT spcname AS "Name",
pg_catalog.pg_get_userbyid(spcowner) AS "Owner",
pg_catalog.pg_tablespace_location(oid) AS "Location"
FROM pg_catalog.pg_tablespace
ORDER BY 1;
**************************

                   List of tablespaces
        Name    |  Owner   |             Location
---------------+----------+----------------------------------
-
    pg_default | postgres |
    pg_global  | postgres |
    tbs_mydb | pguser   | /database/pg10/pg_tbs/tbs_mydb
(3 rows)
```

7.\[illegible]

PostgreSQL[illegible]\[illegible]

```
mydb=> \?
General
    \copyright             show PostgreSQL usage and
distribution terms
    \crosstabview [COLUMNS] execute query and display
results in crosstab
    \errverbose            show most recent error message
at maximum verbosity
    \g [FILE] or ;         execute query (and send results
to file or |pipe)
    \gexec                 execute query, then execute each
value in its result
    \gset [PREFIX]         execute query and store results
in psql variables
    \gx [FILE]             as \g, but forces expanded
```

```
output mode
    \q                      quit psql
    \watch [SEC]            execute query every SEC seconds

Help
    \? [commands]           show help on backslash commands
    \? options              show help on psql command-line
options
    \? variables            show help on special variables
    \h [NAME]               help on syntax of SQL commands,
* for all commands
```

\□□

8.□□□HELP□□

psql□HELP□□□□□□□□□□□\h□□SQL□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□\h CREATE TABLESPACE□□□□□□□□□□□□□□

```
postgres=# \h CREATE TABLESPACE
Command:     CREATE TABLESPACE
Description: define a new tablespace
Syntax:
CREATE TABLESPACE tablespace_name
    [ OWNER { new_owner | CURRENT_USER | SESSION_USER } ]
    LOCATION 'directory'
    [ WITH ( tablespace_option = value [, ... ] ) ]
```

\h□□□□□□□□□SQL□□□□□□□□□SQL□□□□□□□□SQL□□□□□□□□□□□□□

2.2.3 psql[illegible]

psql[illegible]COPY[illegible]\copy[illegible]

1[illegible]COPY[illegible]SQL[illegible]\copy[illegible]

2[illegible]COPY[illegible]SUPERUSER[illegible]stdin[illegible]stdout[illegible]\copy[illegible]SUPERUSER[illegible]

3[illegible]COPY[illegible]\copy[illegible]psql[illegible]

4[illegible]COPY[illegible]\copy[illegible]

1.[illegible]COPY[illegible]

[illegible]COPY[illegible]mydb[illegible]test_copy[illegible]

```
mydb=> CREATE TABLE test_copy(id int4,name text);
CREATE TABLE
```

□□□□□□□test_copy_in.txt□□□□□□□TAB □□□□□□□□□□□□□□□□□□□□□□□□□□□□test-copy-in.txt□□□□□□□

```
[pg10@pghost1 script]$ cat test_copy_in.txt
1       a
2       b
3       c
```

□□□postgres□□□□mydb□□□□ test_copy_in.txt□□□□□□□□□test_copy□□□□□ □□□□□□□

```
[pg10@pghost1 script]$ psql mydb postgres
psql (10.0)
Type "help" for help.

mydb=# COPY pguser.test_copy FROM
'/home/postgres/script/test_copy_in.txt';
COPY 3
mydb=# SELECT * FROM pguser.test_copy ;
    id | name
-------+------
     1 | a
     2 | b
     3 | c
(3 rows)
```

□□□□□□□□pguser□□□□□□□□□□□□□□

```
[pg10@pghost1 script]$ psql mydb pguser
psql (10.0)
```

```
Type "help" for help.

mydb=> COPY test_copy FROM
'/home/postgres/script/test_copy_in.txt';
ERROR:  must be superuser to COPY to or from a file
HINT:  Anyone can COPY to stdout or from stdin. psql's \copy
command also works for anyone.
```

□□□□□□□□□COPY□□□□□□□□□□□□□□\copy□□□□□□□□□□□□□□□□□□□COPY□□□□test_copy□□□□□□□□□□□□□□postgres□□□□□mydb□□□□□□□

```
[pg10@pghost1 script]$ psql mydb postgres
psql (10.0)
Type "help" for help.
mydb=# COPY pguser.test_copy TO
'/home/postgres/test_copy.txt';
COPY 3
```

□□test_copy.txt□□□□□□□□

```
[postgres@pghost1 ~]$ cat test_copy.txt
1       a
2       b
3       c
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ psql mydb pguser
psql (10.0)
```

```
Type "help" for help.

mydb=> COPY test_copy TO stdout;
1       a
2       b
3       c
```

[illegible]

[illegible]DBA[illegible]csv[illegible]

```
[postgres@pghost1 ~]$ psql mydb postgres
psql (10.0)
Type "help" for help.

mydb=# COPY pguser.test_copy TO
'/home/postgres/test_copy.csv' WITH csv header;
COPY 4
```

[illegible]with csv header[illegible]csv[illegible]csv[illegible]office excel[illegible]test_copy[illegible]ID[illegible]1[illegible]

```
mydb=# COPY (SELECT * FROM pguser.test_copy WHERE id=1) TO
'/home/postgres/1.txt';
COPY 1
mydb=# \q
[postgres@pghost1 ~]$ cat 1.txt
1       a
```

□□COPY□□□□□□□□□□
https://www.postgresql.org/docs/10/static/sql-copy.html □

2.□□\copy□□□□□□□□□

COPY□□□□□□□□□□□□□□□□□□□□□□□\copy□□□□psql□□□□□□□□□□□□□□□□□\copy□□□□□□□□□□□□□□□□pghost2□□□□□□□□pguser□□□□pghost1□□□□mydb□□□□□□\copy□□□□□□test_copy□□□□□□□□

```
[postgres@pghost2 ~]$ psql -h 192.168.28.74 -p 1921 mydb
pguser
Password for user pguser:
psql (10.0)
Type "help" for help.

mydb=> \copy test_copy to '/home/postgres/test_copy.txt';
COPY 3
```

□□test_copy.txt□□□□□□□□□□□□□□

```
[postgres@pghost2 ~]$ cat test_copy.txt
1       a
2       b
3       c
```

\copy□□□□□□□copy□□□□□□□□□test_copy_in.txt□□□□□□□□

```
[postgres@pghost2 ~]$ cat test_copy_in.txt
4       d
```

□□\copy□□□□□□test_copy_in.txt□□□□□□□□

```
[postgres@pghost2 ~]$ psql -h 192.168.28.74 -p 1921 mydb
pguser
Password for user pguser:
psql (10.0)
Type "help" for help.

mydb=> \copy test_copy from
'/home/postgres/test_copy_in.txt';
COPY 1
mydb=> SELECT * FROM test_copy WHERE id=4;
    id | name
-------+------
     4 | d
(1 row)
```

□□□□□□□□□□□□□□□□□□□□□□□□□□\copy□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□COPY□□□□□□□□

2.2.4 psql□□□□□□□□

psql□□□□□□□□□□

```
psql [option...] [dbname [username]]
```

□□dbname□□□□□□□□□□username□□□□□□□□□□□option□□□□□□□□□□□□□□□□□□□□

1.-A□□□□□□□□□

psql□□SQL□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ psql -c "SELECT * FROM user_ini WHERE
id=1" mydb pguser
    id | user_id | user_name |          create_time
-------+---------+-----------+-------------------------------
-
     1 |  186536 | KTU89H    | 2017-08-05 15:59:25.359148+08
(1 row)
--□□□□□□□□□□□□
```

□□□□□□□□□□□□□□psql□□-A□□□□□□□

```
[postgres@pghost1 ~]$ psql -A -c "SELECT * FROM user_ini
WHERE id=1" mydb pguser
id|user_id|user_name|create_time
1|186536|KTU89H|2017-08-05 15:59:25.359148+08
(1 row)      --□□□□□□□□□□□
```

[illegible]-A[illegible]-t[illegible]

2.-t[illegible]

[illegible]psql[illegible]-t[illegible]-t[illegible]

```
[postgres@pghost1 ~]$ psql -t -c "SELECT * FROM user_ini
WHERE id=1" mydb pguser
    1 |  186536 | KTU89H    | 2017-08-05 15:59:25.359148+08
--[illegible]
```

[illegible]-t[illegible]-A[illegible]

```
[postgres@pghost1 ~]$ psql -At -c "SELECT * FROM user_ini
WHERE id=1" mydb pguser
1|186536|KTU89H|2017-08-05 15:59:25.359148+08
```

[illegible]shell[illegible]

```
[postgres@pghost1 ~]$ psql -At -c "SELECT user_name FROM
user_ini WHERE id=1" mydb pguser
KTU89H
```

3.-q[illegible]

[illegible]psql[illegible]SQL[illegible]-q[illegible]
test_q.sql[illegible]SQL[illegible]

```
DROP TABLE if exists test_q;
CREATE TABLE test_q(id int4);
TRUNCATE TABLE test_q;
INSERT INTO test_q values (1);
INSERT INTO test_q values (2);
```

[illegible]test_q.sql[illegible]

```
[postgres@pghost1 ~]$ psql mydb pguser -f test_q.sql
DROP TABLE
CREATE TABLE
TRUNCATE TABLE
INSERT 0 1
INSERT 0 1
```

[illegible]test_q.sql[illegible]-q[illegible]

```
[postgres@pghost1 ~]$ psql -q mydb pguser -f test_q.sql
--[illegible]
```

-q[illegible]-c[illegible]-f[illegible]

2.2.5 psql□□sql□□

psql□-c□□□□□□□□□□□□□psql□□□□□□SQL□□□□□□□□

```
[postgres@pghost1 ~]$ psql  -c "SELECT current_user;"
current_user
--------------
postgres
(1 row)
```

-c□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□psql□□-At□□□□□□□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ psql -At -c "SELECT current_user;"
postgres
```

□□□□□□□□□□□□□□□□psql□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□test_2.sql□

```
CREATE TABLE test_2(id int4);
INSERT INTO test_2 VALUES (1);
INSERT INTO test_2 VALUES (2);
INSERT INTO test_2 VALUES (3);
```

□□-f□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ psql mydb pguser -f script/test_2.sql
CREATE TABLE
INSERT 0 1
INSERT 0 1
INSERT 0 1
```

□□□□□□□□□□□□□□□□□□□□SQL□□□□□

□□ psql□-single-transaction□-1□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□SQL□□□□□□□□□□□SQL□□□

2.2.6 psql□□□□□□□SQL

□□□□psql□□□□□□□□SQL□□□□□□SQL□

```
SELECT * FROM table_name WHERE column_name = □□□
```

□□□□□□□□□□□□□

1..\set□□□□□□□□□

\set□□□□□□□□□□□□□□□□□name□□□□□□□value□□□□□□□□□□value□□□□□□□

```
\set name value
```

test_copy□□□□□□□□□□□v_id□□2□□□id□□□2□□□□□□□□□

```
mydb=> \set v_id 2
mydb=> SELECT * FROM test_copy WHERE id=:v_id;
    id | name
-------+------
     2 | b
(1 row)
```

□□□□□□□□□□□□□□\set□□□□□□□□□□□□□□□□

```
mydb=> \set v_id
```

□□\set□□□□□□□□□□□□□□□□□□□pgbench□□□□□□□□□\set□□□□□□□□□

2.psql□-v□□□□□□

□□□□□□□□psql□-v□□□□□□□□□□□select_1.sql□□□□□□□□□□□□

```
SELECT * FROM test_3 WHERE id=:v_id;
```

□□psql□-v□□□□□□□□□□select_1.sql□□□□
□□

```
[postgres@pghost1 ~]$ psql -v v_id=1 mydb pguser -f
select_1.sql
    id | name
-------+------
       1 | a
(1 row)
```

□□□□□□v_id□□1□

2.2.7 □□psql□□□□□□□□

□□□□□□□□□□□□□□□□□□□□DBA□□□psql□□□

1.□□□□□□□□□□□□□

□□□□.psqlrc□□□□□psql□□□-X□□□psql□□□□□□□□□~/.psqlrc□□SQL□□□□□

```
SELECT pid,usename,datname,query,client_addr
FROM pg_stat_activity
```

```
WHERE pid <> pg_backend_pid() AND state='active' ORDER BY
query;
```

---

pg_stat_activity□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□pid□□□□□usename□□□□□□□□□datname□□□□□□□query□□□□□□□□□SQL□□□state□□active□query□□□□□□□□□SQL□client_addr□□□□□□□IP□state□□□□□□□□□□□□

·active□□□□□□□□□SQL□

·idle□□□□□□□□□□□□□□□□□□□□□□□□

·idle in transaction□□□□□□□□□□□□□□□□□□□□SQL□

·idle in transaction□aborted□□□idle in transaction□□□□□□□□□□□□□SQL□□□

□□□□□□□□□□□□□□ https://www.postgresql.org/docs/10/static/monitoring-stats.html#pg-stat-activity-view □

□□□□~/.psqlrc□□□□□□□□□□□□□□□□□□□□□□□□□□\set□□□□□SQL□□□□□□□□□

```
--□□□□□□
\set active_session 'select
pid,usename,datname,query,client_addr from pg_stat_activity
where pid <> pg_backend_pid() and state=\'active\' order by
query;'
```

□□□□□□□□□□□□□active_session□□□□□□□
□□□□□□□□□□□

```
postgres=# :active_session
      pid | usename | datname |                  query
| client_addr
----------+--------+--------+----------------------------
-------------------------+
    14351 | pguser  | mydb    | update test_per1 set
create_time=now() WHERE id=$1; |
    14352 | pguser  | mydb    | update test_per1 set
create_time=now() WHERE id=$1; |
    14353 | pguser  | mydb    | update test_per1 set
create_time=now() WHERE id=$1; |
    14354 | pguser  | mydb    | update test_per1 set
create_time=now() WHERE id=$1; |
    14355 | pguser  | mydb    | update test_per1 set
create_time=now() WHERE id=$1; |
(5 rows)
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
SQL□□□□□□active_session□□□□□□□□□□□□□

2.□□□□□□□□□□□□□

PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□SQL□□□□□

```
SELECT
pid,usename,datname,query,client_addr,wait_event_type,wait_e
vent
FROM pg_stat_activity
WHERE pid <> pg_backend_pid() AND wait_event is not null
ORDER BY wait_event_type;
```

□□□□□\set□□□□□□□□□□□~/.psqlrc□□□□□
\set□□□SQL□□□□□□□□□□□□□□□

```
--□□□□□□□□□
\set wait_event 'select
pid,usename,datname,query,client_addr,wait_event_type,wait_e
vent from pg_stat_activity where pid <> pg_backend_pid() and
wait_event is not null order by wait_event_type;
```

□□□□□□□□□□□□□wait_event□□□□□□□□□□
□□□□□□□□

```
postgres=# :wait_event
    pid  | usename  | datname | query | client_addr |
wait_event_type |     wait_event
---------+----------+--------+-------+-----------+---------
------+-------------------
    2652 |          |        |      |            | Activity
| AutoVacuumMain
    2655 | postgres  |        |        |             |
Activity        | LogicalLauncherMain
    2650 |           |         |       |             |
Activity        | BgWriter Hibernate
    2649 |          |        |       |             |
Activity       | CheckpointerMain
    2651 |          |        |       |            | Activity
| WalWriterMain
(5 rows)
```

[illegible]

```
--[illegible]
\set connections 'select
datname,usename,client_addr,count(*) from pg_stat_activity
where pid <> pg_backend_pid() group by 1,2,3 order by 1,2,4
desc;'
```

[illegible]\set[illegible]Zabbix[illegible]Nagios[illegible]

2.2.8 psql[illegible]

psql[illegible]SQL[illegible]SQL[illegible]psql[illegible]

1.\timing[illegible]SQL[illegible]

\timing[illegible]SQL[illegible]

```
mydb=> \timing
Timing is on.
mydb=> SELECT count(*) FROM user_ini;
```

```
     count
---------
    1000000
(1 row)

Time: 47.114 ms
```

□□□□count□□□□□□□□47.114□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□\timing□□□□□□□□□□□□

```
mydb=> \timing
Timing is off.
```

2.\watch□□□□□□SQL

\watch□□□□□□□□□□□□□□□□SQL□□□□□SQL□□□□□□□□□□□□□□

```
\watch [ seconds ]
```

seconds□□□□□□□□□□□□□□□□□□□□□2□□□□□□□□□□□□□now□□□□□□□□□□□

```
mydb=> SELECT now();
              now
-------------------------------
 2017-08-14 11:20:02.157567+08
(1 row)
```

```
mydb=> \watch 1
Mon 14 Aug 2017 11:20:04 AM CST (every 1s)

              now
-------------------------------
 2017-08-14 11:20:04.299584+08
(1 row)

Mon 14 Aug 2017 11:20:05 AM CST (every 1s)

              now
-------------------------------
 2017-08-14 11:20:05.300991+08
```

[illegible]now[illegible]

3.Tab[illegible]

psql[illegible]Tab[illegible]SQL[illegible]mydb[illegible]test[illegible]test[illegible]Tab[illegible]psql[illegible]test[illegible]

```
mydb=> SELECT * FROM test_
test_1     test_2     test_copy  test_per1
mydb=> SELECT * FROM test_
```

DDL[illegible]Tab[illegible]

```
mydb=> ALTER TABLE test_1 DROP CO
COLUMN      CONSTRAINT
mydb=> ALTER TABLE test_1 DROP CO
```

4.□□□□□□□□□□SQL□□

psql□□□□□□□□□□SQL□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ psql mydb pguser
psql (10.0)
Type "help" for help.

mydb=> SELECT count(*) FROM pg_stat_activity ; --□□□□□□□□□□
□□□
```

□□psql□□□□□□□□□□SQL□□□□□□□□PostgreSQL□□□□readline□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□--without-readline□□□□readline□□□□□□

5.psql□□□□□□

□□□□□□□psql□□□□□□□“postgres=#”□□□□□□□□□□□

```
[postgres@pghost1 ~]$ psql
psql (10.0)
Type "help" for help.

postgres=#
```

[illegible]psql[illegible]psql[illegible]

·%M[illegible]psql[illegible]-h[illegible]Unix[illegible][local][illegible]

·%>[illegible]

·%n[illegible]SET SESSION AUTHORIZATION[illegible]

·%/[illegible]

·%#[illegible]“#”[illegible]“>”[illegible]SET SESSION AUTHORIZATION[illegible]

·%p[illegible]

·%R[illegible]PROMPT1[illegible]“=”[illegible]“[illegible]”[illegible]

[illegible]psql[illegible]prompt1[illegible]

```
[postgres@pghost1 ~]$ psql
psql (10.0)
Type "help" for help.
```

```
postgres=# \echo :PROMPT1
%/%R%#
```

□□□\echo□□□□□□□□PROMPT1□□□□□□□□□□PROMPT1□□□psql□□□□□□□□□□□□□□□□□□□□□□%/%R%#□□□□□□□□□□□□□□□□□%/□□□□□□□□postgres□%R□□□□□“=”□%#□□□□“#”□□□□□%M□□□□□□□□□□□□□□Unix□□□□□□□□□PROMPT1□□%M%R%#□□□□□□

```
[postgres@pghost1 ~]$ psql
psql (10.0)
Type "help" for help.

postgres=# \set PROMPT1 '%M%R%#'
[local]=#
```

□□□psql□□□□□[local]□□□□□pghost2□□□□□□□pghost1□□□□□□□□□PROMPT1□□□□“%M%R%#”□□□□□□

```
[postgres@pghost2 ~]$ psql -h 192.168.28.74 mydb pguser -p
1921
Password for user pguser:
psql (10.0)
Type "help" for help.

mydb=> \set PROMPT1  '%M%R%#'
192.168.28.74=>
```

[illegible]PROMPT1[illegible]“%M%R%#”[illegible]192.168.28.74[illegible]IP[illegible]psql[illegible]-h[illegible]

```
[postgres@pghost2 ~]$ psql -h 192.168.28.74 mydb pguser -p
1921
Password for user pguser:
psql (10.0)
Type "help" for help.

mydb=> \set PROMPT1  '%/@%M:%>%R%#'
mydb@192.168.28.74:1921=>
```

[illegible]PROMPT1[illegible]“%/@%M[illegible]%>%R%#”[illegible]“%>”[illegible]PROMPT1[illegible]PROMPT1[illegible].psqlrc[illegible].psqlrc[illegible]2.2.7[illegible].psqlrc[illegible]

```
[postgres@pghost2 ~]$ touch .psqlrc
 [postgres@pghost2 ~]$ vim .psqlrc
\set PROMPT1  '%/@%M:%>%R%#'
```

[illegible]

```
[postgres@pghost2 ~]$ psql -h 192.168.28.74 mydb pguser -p
1921
Password for user pguser:
psql (10.0)
```

```
Type "help" for help.

mydb@192.168.28.74:1921=>
```

□□□□□□□□□□□□PROMPT1□□□.psqlrc□□□psql□□□□□□□□□.psqlrc□□□□□□□□□□□

□□ psql□□□□□□□□□PROMPT1□PROMPT2□PROMPT3□PROMPT1□□□psql□□□□□□□□□□□□□□□□□□□□□□□□□□PROMPT2□□PROMPT2□□□□□□□PROMPT1□□□PROMPT3□□□□□□SQLCOPYFROMSTDIN□□□□□□□□□□□□□□□□□□□□□□□□

2.3 □□□□

□□□□□PostgreSQL□□□□□□□pgAdmin 4□psql□□□□□□□□□□□□□psql□□□□□□□□□□□□□□□□□□□□□□□□□□□□pgAdmin 4□□□□□□□□□□□□□psql□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□psql□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□3□ □□□□

□□□□□PostgreSQL□□□□□□□PostgreSQL□□/□□json/jsonb□□□

3.1 [illegible]

PostgreSQL[illegible]serial[illegible]

3.1.1 [illegible]

PostgreSQL[illegible]3-1[illegible]

[illegible]3-1 [illegible]

类型名称	存储长度	描述	范围
smallint	2 字节	小范围整数类型	32 768 到 +32 767
integer	4 字节	整数类型	–2 147 483 648 到 +2 147 483 647
bigint	8 字节	大范围整数类型	–9 223 372 036 854 775 808 到 +9 223 372 036 854 775 807
decimal	可变	用户指定精度	小数点前 131 072 位；小数点后 16 383 位
numeric	可变	用户指定精度	小数点前 131 072 位；小数点后 16 383 位
real	4 字节	变长，不精确	6 位十进制精度
double precision	8 字节	变长，不精确	15 位十进制精度
smallserial	2 字节	smallint 自增序列	1 到 32 767
serial	4 字节	integer 自增序列	1 到 2 147 483 647
bigserial	8 字节	bigint 自增序列	1 到 9 223 372 036 854 775 807

smallint[illegible]integer[illegible]bigint[illegible]smallint[illegible]2[illegible]int2[illegible]integer[illegible]4[illegible]smallint[illegible]int4[illegible]bigint[illegible]8[illegible]

□□□□□□integer□□□□□□□□□□int8□□□□□□□□□□□
□□□□□□integer□□□□□□integer□□□□□□□□□□□□
bigint□□□□□□□integer□□□□□□□□□

```
mydb=> CREATE TABLE test_integer (id1 integer,id2 int4) ;
CREATE TABLE
```

decimal□numeric□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□numeric
□□□□□□□□

```
NUMERIC(precision, scale)
```

precision□□numeric□□□□□□□□□scale□□□
□□□□□□□□□□□18.222□precision□5□□scale□
3□precision□□□□□□□scale□□□0□□□□□□
numeric□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□numeric□□□□□

real□double precision□□□□□□□□□real□□
4□□□double precision□□8□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□

smallserial□serial□bigserial□□□□□□
serial□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□test_serial□□id□□□serial□□□

```
mydb=> CREATE TABLE test_serial (id serial,flag text);
CREATE TABLE
```

[illegible]test_serial[illegible]

```
mydb=> \d test_serial
                             Table "pguser.test_serial"
    Column |  Type   | Collation | Nullable |
Default
-----------+---------+-----------+----------+--------------
----------------------
    id     | integer |           | not null |
nextval('test_serial_id_seq'::regclass)
    flag   | text    |           |          |
```

[illegible]id[illegible]test_serial_id_seq[illegible]serial[illegible]

```
mydb=> INSERT INTO test_serial(flag) VALUES ('a');
INSERT 0 1
mydb=> INSERT INTO test_serial(flag) VALUES ('b');
INSERT 0 1
mydb=> INSERT INTO test_serial(flag) VALUES ('c');
INSERT 0 1
mydb=> SELECT * FROM test_serial;
    id | flag
-------+------
     1 | a
     2 | b
     3 | c
(3 rows)
```

3.1.2 [illegible]

PostgreSQL[illegible]

```
mydb=> SELECT 1+2,2*3,4/2,8%3;
    ?column? | ?column? | ?column? | ?column?
-------------+----------+----------+----------
           3 |        6 |        2 |        2
```

[illegible]

```
mydb=> SELECT mod(8,3);
 mod
-----
 2
(1 row)
```

[illegible]

```
mydb=> SELECT round(10.2),round(10.9);
    round | round
----------+-------
       10 |    11
(1 row)
```

[illegible]

```
mydb=> SELECT ceil(3.6),ceil(-3.6);
    ceil | ceil
---------+------
       4 |   -3
(1 row)
```

[illegible]

```
mydb=> SELECT floor(3.6), floor(-3.6);
    floor | floor
----------+-------
        3 |    -4
(1 row)
```

3.2 □□□□

□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□
□□□

3.2.1 □□□□□□

PostgreSQL□□□□□□□□□3-2□□□

□3-2 □□□□□□□□

字符类型名称	描　　述
character varying(n), varchar(n)	变长，字符最大数有限制
character(n), char(n)	定长，字符数没达到最大值则使用空白填充
text	变长，无长度限制

character varying□n□□□□□□□□□□□□□n□□□□□□□□□□□□□□□□□□□n□□□□□□□□□□□□□□□n□□character varying□n□□□□□□□□□□□□□character□n□□□□□□□□□□□□□□□□□□□□n□□□□□□□□□□□□□□□n□□

```
mydb=> CREATE TABLE test_char(col1 varchar (4),col2
character(4));
CREATE TABLE
mydb=> INSERT INTO test_char(col1,col2) VALUES ('a','a');
INSERT 0 1
```

□test_char□□□col1□□□character varying□4□□col2□□□character□4□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT char_length(col1),char_length(col2) FROM
test_char ;
    char_length | char_length
---------------+-------------
              1 |           1
(1 row)
```

char_length□string□□□□□□□□□□□□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT octet_length(col1),octet_length(col2) FROM
test_char ;
    octet_length | octet_length
----------------+--------------
               1 |            4
(1 row)
```

octet_length□string□□□□□□□□□□□□□col2□□□□□4□□□□□□□col2□□□□□character□□□

□□□□□□character varying□n□□□□□□□□□□□□□□□□□□□□□□□□□character□n□□□□□□□□□□□□□character□1□□□

text[illegible]character varying[illegible]

[illegible] PostgreSQL[illegible]1GB[illegible]character varying[illegible]text[illegible]1GB[illegible]character[illegible]n[illegible]

3.2.2 [illegible]

PostgreSQL[illegible]

[illegible]

```
mydb=> SELECT char_length('abcd');
 char_length
-------------
           4
(1 row)
```

[illegible]

```
mydb=> SELECT octet_length('abcd');
 octet_length
--------------
            4
(1 row)
```

□□□□□□□□□□□□□□□□□

```
mydb=> SELECT position('a' in 'abcd');
 position
----------
        1
(1 row)
```

□□□□□□□□□□□□□□□

```
mydb=> SELECT substring('francs' from 3 for 4);
 substring
-----------
 ancs
(1 row)
```

□□□□□□split_part□□□□□□□

```
split_part(string text, delimiter text, field int)
```

□□delimiter□□□□□□□□string□□□□□□□□□□□□1□□□□□□□□

```
mydb=> SELECT split_part('abc@def1@nb','@',2);
 split_part
------------
 def1
(1 row)
```

3.3 □□/□□□□

PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□

3.3.1 □□/□□□□□□

PostgreSQL□□□□□□□□□□□□3-3□□□

□3-3 □□□□□□□□□□□

字符类型名称	存储长度	描述
timestamp [(p)] [without time zone]	8 字节	包括日期和时间，不带时区，简写成 timestamp
timestamp [(p)] with time zone	8 字节	包括日期和时间，带时区，简写成 timestamptz
date	4 字节	日期，但不包含一天中的时间
time [(p)] [without time zone]	8 字节	一天中的时间，不包含日期，不带时区
time [(p)] with time zone	12 字节	一天中的时间，不包含日期，带时区
interval [fields] [(p)]	16 字节	时间间隔

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□now□□□□□now□□□□□□□□□□□□□□□□□timestamp[□p□]with time zone□□□□□□

```
mydb=> SELECT now();
              now
-------------------------------
 2017-07-29 09:44:25.493425+08
(1 row)
```

[illegible]SQL[illegible]timestamp without time zone[illegible]

```
mydb=> SELECT now()::timestamp without time zone;
              now
----------------------------
    2017-07-29 09:44:55.804403
(1 row)
```

[illegible]date[illegible]

```
mydb=> SELECT now()::date;
   now
------------
 2017-07-29
(1 row)
```

[illegible]time without time zone[illegible]

```
mydb=> SELECT now()::time without time zone;
    now
-----------------
09:45:49.390428
(1 row)
```

[illegible]time with time zone[illegible]

```
mydb=> SELECT now()::
time with time zone;
```

```
        now
-------------------
 09:45:57.13139+08
(1 row)
```

interval□□□□□□□□□□□□□□□hour□day□month□year□□□□□□□

```
mydb=> SELECT now(),now()+interval'1 day';
                 now               |           ?column?
-----------------------------------+-------------------------------
 2017-07-29 09:47:26.026418+08 | 2017-07-30 09:47:26.026418+08
(1 row)
```

□□p□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□6□□□□□□□□□□0□

```
mydb=> SELECT now(), now()::timestamp(0);
                 now               |         now
-----------------------------------+---------------------
 2017-07-29 09:59:42.688445+08 | 2017-07-29 09:59:43
(1 row)
```

3.3.2 □□/□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□

```
mydb=> SELECT date '2017-07-29' +  interval'1 days';
        ?column?
---------------------
    2017-07-30 00:00:00
(1 row)
```

□□□□□□□□□□

```
mydb=> SELECT date '2017-07-29' -  interval'1 hour';
        ?column?
---------------------
    2017-07-28 23:00:00
(1 row)
```

□□□□□□□□□□

```
mydb=> SELECT 100* interval '1 second';
    ?column?
----------
    00:01:40
(1 row)
```

□□□□□□□□□□

```
mydb=> SELECT interval '1 hour' / double precision '3';
    ?column?
----------
    00:20:00
(1 row)
```

3.3.3 [illegible]/[illegible]

[illegible]

[illegible]

```
mydb=> SELECT current_date, current_time;
    current_date |    current_time
-----------------+--------------------
    2017-07-29   | 10:53:10.375374+08
(1 row)
```

[illegible]EXTRACT[illegible]

```
EXTRACT(field FROM source)
```

field[illegible]century[illegible]year[illegible]month[illegible]day[illegible]hour[illegible]minute[illegible]second[illegible]source[illegible]timestamp[illegible]time[illegible]interval[illegible]

```
mydb=> SELECT EXTRACT( year FROM now());
    date_part
-----------
        2017
(1 row)
```

[illegible]timestamp[illegible]

```
mydb=> SELECT EXTRACT( month FROM now()),EXTRACT(day FROM
now());
    date_part | date_part
--------------+-----------
            7 |        29
(1 row)
```

[illegible]

```
mydb=> SELECT EXTRACT( hour FROM now()), extract (minute
FROM now());
    date_part | date_part
--------------+-----------
           11 |        14
```

[illegible]

```
mydb=> SELECT EXTRACT( second FROM now());
    date_part
-----------
    43.031366
(1 row)
```

[illegible]

```
mydb=> SELECT EXTRACT( week FROM now());
    date_part
-----------
        30
(1 row)
```

□□□□□□□□□□□□□□□□

```
mydb=> SELECT EXTRACT( doy FROM now());
    date_part
-----------
        210
(1 row)
```

3.4 □□□□

□□□□□□□PostgreSQL□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□json/jsonb□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□3-4□□□

□3-4 □□□□□□

字符类型名称	存储长度	描述
boolean	1字节	状态为true或false

true□□□□□□□□□TRUE□t□true□y□yes□on□1□false□□□□□□□FALSE□f□false□n□no□off□0□□□□□□□□□□□□□□□□□□□

```
mydb=> CREATE TABLE test_boolean(cola boolean,colb boolean);
CREATE TABLE
mydb=> INSERT INTO test_boolean (cola,colb) VALUES
('true','false');
INSERT 0 1
mydb=> INSERT INTO test_boolean (cola,colb) VALUES ('t','f');
INSERT 0 1
mydb=> INSERT INTO test_boolean (cola,colb) VALUES
('TRUE','FALSE');
INSERT 0 1
mydb=> INSERT INTO test_boolean (cola,colb) VALUES
('yes','no');
INSERT 0 1
mydb=> INSERT INTO test_boolean (cola,colb) VALUES ('y','n');
INSERT 0 1
```

```
mydb=> INSERT INTO test_boolean (cola,colb) VALUES ('1','0');
INSERT 0 1
mydb=> INSERT INTO test_boolean (cola,colb) VALUES
(null,null);
INSERT 0 1
```

□□□test_boolean□□□□□□□□□true□false□□□□□□□□□□□□□□□true□□□□□t□false□□□□□f□□□□□□□NULL□□□□□□□□□□□□

```
mydb=> SELECT * FROM test_boolean ;
    cola | colb
---------+------
    t    | f
    t    | f
    t    | f
    t    | f
    t    | f
    t    | f
         |
(7 rows)
```

3.5 □□□□□□

□□□□IP□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□□IPv4□IPv6□MAC□□□□□□□□□□□□□□□□□□□□□□□□□□□□IP□□□

3.5.1 □□□□□□□□

PostgreSQL□□□□□□□□□□□□□3-5□□□

□3-5 □□□□□□□□□□

字符类型名称	存储长度	描述
cidr	7 或 19 字节	IPv4 和 IPv6 网络
inet	7 或 19 字节	IPv4 和 IPv6 网络
macaddr	6 字节	MAC 地址
macaddr8	8 字节	MAC 地址（EUI-64 格式）

inet□cidr□□□□□□□□□□□□address/y□□□address□□IPv4□IPv6□□□□□y□□□□□□□□□□□y□□□□□□IPv4□□□□□32□□□IPv6□□□□□128□□□□□□□□□□□□

inet□cidr□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT '192.168.2.1000'::inet;
ERROR:  invalid input syntax for type inet: "192.168.2.1000"
LINE 1: select '192.168.2.1000'::inet;
```

inet□cidr□□□□□□□□

1□cidr□□□□□□□□□□□□inet□□□□□□□

```
mydb=> SELECT '192.168.1.100'::cidr;
        cidr
------------------
    192.168.1.100/32
(1 row)

mydb=> SELECT '192.168.1.100/32'::inet;
        inet
---------------
    192.168.1.100
(1 row)

mydb=> SELECT '192.168.0.0/16'::inet;
        inet
----------------
    192.168.0.0/16
(1 row)
```

2□cidr□□□IP□□□□□□□□□□□□inet□□□□□□

```
mydb=> SELECT '192.168.2.0/8'::cidr;
ERROR:  invalid cidr value: "192.168.2.0/8"
LINE 1: select '192.168.2.0/8'::cidr;
DETAIL:  Value has bits set to right of mask.
```

```
mydb=> SELECT '192.168.2.0/8'::inet;
        inet
---------------
    192.168.2.0/8
(1 row)

mydb=> SELECT '192.168.2.0/24'::cidr;
        cidr
----------------
    192.168.2.0/24
(1 row)
```

□□□□□□□□□□cidr□inet□□□□□□□□macaddr□macaddr8□□MAC□□□□□□□□□□

3.5.2 □□□□□□□

PostgreSQL□□□□□□□□□□□□□□□□□□□3-6□□□

□3-6 □□□□□□□□□□□

操 作 符	描 述	举 例
<	小于	inet '192.168.1.5' < inet '192.168.1.6'
<=	小于等于	inet '192.168.1.5' <= inet '192.168.1.5'
=	等于	inet '192.168.1.5' = inet '192.168.1.5'
>=	大于等于	inet '192.168.1.5' >= inet '192.168.1.5'
>	大于	inet '192.168.1.5' > inet '192.168.1.4'
<>	不等于	inet '192.168.1.5' <> inet '192.168.1.4'
<<	被包含	inet '192.168.1.5' << inet '192.168.1/24'
<<=	被包含或等于	inet '192.168.1/24' <<= inet '192.168.1/24'
>>	包含	inet '192.168.1/24' >> inet '192.168.1.5'
>>=	包含或等于	inet '192.168.1/24' >>= inet '192.168.1/24'
&&	包含或被包含	inet '192.168.1/24' && inet '192.168.1.80/28'
~	按位取反	~ inet '192.168.1.6'
&	按位与	inet '192.168.1.6' & inet '0.0.0.255'
|	按位或	inet '192.168.1.6' | inet '0.0.0.255'
+	加	inet '192.168.1.6' + 25
-	减	inet '192.168.1.43' - 36
-	减	inet '192.168.1.43' - inet '192.168.1.19

3.5.3 [illegible]

PostgreSQL[illegible]

[illegible]IP[illegible]

```
mydb=> SELECT host(cidr '192.168.1.0/24');
        host
-------------
    192.168.1.0
(1 row)
```

[illegible]IP[illegible]

```
mydb=> SELECT text(cidr '192.168.1.0/24');
        text
----------------
    192.168.1.0/24
(1 row)
```

[illegible]

```
mydb=> SELECT netmask(cidr '192.168.1.0/24');
        netmask
---------------
    255.255.255.0
```

3.6 [illegible]

PostgreSQL[illegible]
[illegible]

3.6.1 [illegible]

[illegible]
[illegible]“[]”[illegible]

```
CREATE TABLE test_array1 (
    id          integer,
    array_i     integer[],
    array_t     text[]
);
```

[illegible]integer[][illegible]integer[illegible]text[][illegible]
text[illegible]

3.6.2 [illegible]

[illegible]

```
'{ val1 delim val2 delim ... }'
```

□□□□□□□□“{}”□□□delim□□□□□□□□
□□□□□□□□delim□□□□□□□□□□□

```
mydb=> SELECT '{1,2,3}';
    ?column?
----------
    {1,2,3}
(1 row)
```

□□test_array1□□□□□□□□□□□□□□□

```
mydb=> INSERT INTO test_array1(id,array_i,array_t)
VALUES (1,'{1,2,3}','{"a","b","c"}');
INSERT 0 1
```

□□□□□□□□□□□□□□□ARRAY□□□□□□□

```
mydb=> SELECT array[1,2,3];
        array
---------
    {1,2,3}
(1 row)
```

□test_array2□□□□□□□□□□□□□□□□□

```
mydb=> INSERT INTO test_array1(id,array_i,array_t)
    VALUES (2,array[4,5,6],array['d','e','f']);
INSERT 0 1
```

□test_array2□□□□□□□□

```
mydb=> SELECT * FROM test_array1;
    id | array_i | array_t
-------+---------+---------
     1 | {1,2,3} | {a,b,c}
     2 | {4,5,6} | {d,e,f}
(2 rows)
```

3.6.3 □□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT array_i FROM test_array1 WHERE id=1;
    array_i
---------
    {1,2,3}
(1 row)
```

□□□□□□□□□□□□“[]”□□□□□□□□□□□□□□□□□□□□□□□□□□□□1□n□n□□□□□□□□□□□

```
mydb=> SELECT array_i[1],array_t[3] FROM test_array1 WHERE
id=1;
    array_i | array_t
------------+---------
          1 | c
(1 row)
```

3.6.4 □□□□□□□□□□□

PostgreSQL[illegible]
[illegible]array_append[illegible]

```
array_append(anyarray, anyelement)
```

array_append[illegible]

```
mydb=> SELECT array_append(array[1,2,3],4);
    array_append
--------------
    {1,2,3,4}
(1 row)
```

[illegible]||[illegible]

```
mydb=> SELECT array[1,2,3] || 4;
    ?column?
-----------
    {1,2,3,4}
(1 row)
```

[illegible]array_remove[illegible]
array_remove[illegible]

```
array_remove(anyarray, anyelement)
```

array_remove[illegible]

```
mydb=> SELECT array[1,2,2,3],array_remove(array[1,2,2,3],2);
    array    | array_remove
------------+--------------
  {1,2,2,3} | {1,3}
(1 row)
```

[illegible]

```
mydb=> UPDATE test_array1 SET array_i[3]=4 WHERE id=1 ;
UPDATE 1
```

[illegible]

```
mydb=> UPDATE test_array1 SET array_i=array[7,8,9] WHERE
id=1;
UPDATE 1
```

3.6.5 [illegible]

PostgreSQL[illegible]3-7[illegible]

[illegible]3-7 [illegible]

操 作 符	描 述	举 例	结 果
=	等于	ARRAY[1.1,2.1,3.1]::int[] = ARRAY[1,2,3]	t
<>	不等于	ARRAY[1,2,3] <> ARRAY[1,2,4]	t
<	小于	ARRAY[1,2,3] < ARRAY[1,2,4]	t
>	大于	ARRAY[1,4,3] > ARRAY[1,2,4]	t
<=	小于等于	ARRAY[1,2,3] <= ARRAY[1,2,3]	t
>=	大于等于	ARRAY[1,4,3] >= ARRAY[1,4,3]	t
@>	包含	ARRAY[1,4,3] @> ARRAY[3,1]	t
<@	被包含	ARRAY[2,7] <@ ARRAY[1,7,4,2,6]	t
&&	重叠（具有公共元素）	ARRAY[1,4,3] && ARRAY[2,1]	t
||	数组和数组串接	ARRAY[1,2,3] || ARRAY[4,5,6]	{1,2,3,4,5,6}
||	数组和数组串接	ARRAY[1,2,3] || ARRAY[[4,5,6],[7,8,9]]	{{1,2,3},{4,5,6},{7,8,9}}
||	元素和数组串接	3 || ARRAY[4,5,6]	{3,4,5,6}
||	数组和元素串接	ARRAY[4,5,6] || 7	{4,5,6,7}

3.6.6 □□□□

PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT
array_append(array[1,2],3),array_remove(array[1,2],2);
    array_append | array_remove
-----------------+--------------
    {1,2,3}      | {1}
(1 row)
```

□□□□□□□□□□□□

```
mydb=> SELECT array_ndims(array[1,2]);
    array_ndims
-------------
```

```
        1
(1 row)
```

[illegible]

```
mydb=> SELECT array_length(array[1,2],1);
    array_length
--------------
        2
(1 row)
```

[illegible]

```
mydb=> SELECT array_position(array['a','b','c','d'],'d');
    array_position
----------------
        4
(1 row)
```

[illegible]array_replace[illegible]

```
array_replace(anyarray, anyelement, anyelement)
```

[illegible]anyarray[illegible]anyelement[illegible]

```
mydb=> SELECT array_replace(array[1,2,5,4],5,10);
    array_replace
---------------
```

```
    {1,2,10,4}
(1 row)
```

□□□□□□□□□□□□□□□□array_to_string□□□□□□□□

```
atray_to_string(anyarray, text [, text])
```

□□□□□□□□text□□□□text□□□□□□□□□□text□□□□□NULL□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT array_to_string(array[1,2,null,3],',','10');
    array_to_string
-----------------
    1,2,10,3
(1 row)
```

3.7 [illegible]

[illegible]

3.7.1 [illegible]

PostgreSQL[illegible]

·int4range—integer[illegible]

·int8range—bigint[illegible]

·numrange—numeric[illegible]

·tsrange—[illegible]timestamp[illegible]

·tstzrange—[illegible]timestamp[illegible]

·daterange—date[illegible]

[illegible]CREATE TYPE[illegible]integer[illegible]

```
mydb=> SELECT int4range(1,5);
    int4range
-----------
    [1,5)
(1 row)
```

□□□□1□5□□□□□□date□□□□□□□□□

```
mydb=> SELECT daterange('2017-07-01','2017-07-30');
        daterange
-------------------------
    [2017-07-01,2017-07-30)
```

3.7.2 □□□□□□

□□□□□□□□□□□□□□□□□□□“[”□□□□□□□□□□“□”□□□□□□□□□□“]”□□□□□□□□□□“□”□□

```
(lower-bound,upper-bound)
(lower-bound,upper-bound]
[lower-bound,upper-bound)
[lower-bound,upper-bound]
empty
```

□□empty□□□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT int4range(4,7);
    int4range
-----------
    [4,7)
(1 row)
```

[illegible]4[illegible]5[illegible]6[illegible]7[illegible]

```
mydb=> SELECT int4range(1,3);
    int4range
-----------
    [1,3)
(1 row)
```

[illegible]“]”[illegible]

```
mydb=> SELECT int4range(1,3,'[]');
    int4range
-----------
    [1,4)
(1 row)
```

[illegible]“]”[illegible]“[illegible]”[illegible]

3.7.3 [illegible]

[illegible]

[illegible]

```
mydb=> SELECT int4range(4,7) @> 4;
    ?column?
----------
    t
(1 row)
```

[illegible]

```
mydb=> SELECT int4range(4,7)@>int4range(4,6);
    ?column?
----------
    t
(1 row)
```

[illegible]

```
mydb=> SELECT int4range(4,7)=int4range(4,6,'[]');
    ?column?
----------
    t
(1 row)
```

[illegible]“@>”[illegible]

3.7.4 [illegible]

□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT lower(int4range(1,10));
    lower
-------
    1
(1 row)
```

□□□□□□□□□□□

```
mydb=> SELECT upper(int4range(1,10));
    upper
-------
    10
(1 row)
```

□□□□□□□□□□□□

```
mydb=> SELECT isempty(int4range(1,10));
    isempty
---------
    f
(1 row)
```

3.7.5 □□□□□□□□□

□□□□□□□□□□GiST□□□GiST□□□□□□□□□“=”“&&”“<@”“@>”“<<”“>>”“-|-”“&<”“&>”□□GiST□□□□□□□□□

```
CREATE INDEX idx_ip_address_range ON ip_address USING gist (
ip_range);
```

3.8 json/jsonb□□

PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□json□JavaScript Object Notation□□json□□□□□□□□□□□□□□□□□□□□json□□□json□jsonb□□□json□jsonb□□□□□□□□□□jsonb□□□□□□□□□□□□□

3.8.1 json□□□□

PostgreSQL□□9.2□□□□□□□□json□□□□□□□□□□□□□□□□PostgreSQL□json□□□□□□□□□□□□□□□□json□□□□□□□□□□□□□□□□□□□□json□□□□□□□□

```
mydb=> SELECT '{"a":1,"b":2}'::json;
        json
---------------
    {"a":1,"b":2}
```

□□□□□□□□json□□□□□□□□□□□□□□□□□□□

```
mydb=> CREATE TABLE test_json1 (id serial primary key,name
json);
CREATE TABLE
```

□□□□□□□□□name□json□□□□□□□□□□□□□□□

```
mydb=> INSERT INTO test_json1 (name)
VALUES ('{"col1":1,"col2":"francs","col3":"male"}');
INSERT 0 1

mydb=> INSERT INTO test_json1 (name)
VALUES ('{"col1":2,"col2":"fp","col3":"female"}');
INSERT 0 1
```

□□□test_json1□□□□□□□□

```
mydb=> SELECT * FROM test_json1;
    id |                   name
-------+----------------------------------------
     1 | {"col1":1,"col2":"francs","col3":"male"}
     2 | {"col1":2,"col2":"fp","col3":"female"}
```

3.8.2 □□json□□

□□“->”□□□□□□□json□□□□□□□□□□□

```
mydb=> SELECT  name -> 'col2' FROM test_json1 WHERE id=1;
    ?column?
----------
    "francs"
(1 row)
```

□□□□□□□□□□json□□□□□□□□“->>”□□□□□□□□□

```
mydb=> SELECT  name ->> 'col2' FROM test_json1 WHERE id=1;
    ?column?
----------
    francs
(1 row)
```

3.8.3 jsonb□json□□

PostgreSQL□□□□JSON□□□□□json□jsonb□□□□□□□□□□□□□□□□□□□□□□□□□□□□□json□□□□□□□□jsonb□□□□□□□□□□□□□□□□□□□□□□□json□□□□□□□□□□□□□□json□□□□□□□□□□□□□□□□□□□□□□□json□□□□□□□□□□□jsonb□□□□□□□□□□□□□□□□□□□□jsonb□□□□□□□□□□□□□□□json□□□jsonb□□□□□□jsonb□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□json□jsonb□□□□□□□□□□□□□□□□jsonb□□□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT '{"bar": "baz", "balance": 7.77,
"active":false}'::jsonb;
        jsonb
--------------------------------------------------
    {"bar": "baz", "active": false, "balance": 7.77}
(1 row)
```

□json□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT '{"bar": "baz", "balance": 7.77,
"active":false}'::json;
        json
-------------------------------------------------
    {"bar": "baz", "balance": 7.77, "active":false}
(1 row)
```

□□□jsonb□□□□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT ' {"id":1,    "name":"francs"}'::jsonb;
        jsonb
-----------------------------
    {"id": 1, "name": "francs"}
(1 row)
```

□□□id□□name□□□□□□□□□□□□□□□□□□□□□
□json□□□□□□□□□□□□□□□□□

```
mydb=> SELECT ' {"id":1,    "name":"francs"}'::json;
            json
-------------------------------
        {"id":1,    "name":"francs"}
(1 row)
```

□□□jsonb□□□□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT ' {"id":1,
"name":"francs",
"remark":"a good guy!",
"name":"test"
}'::jsonb;
                jsonb
----------------------------------------------------
```

```
    {"id": 1, "name": "test", "remark": "a good guy!"}
(1 row)
```

[illegible]name[illegible]name[illegible]json[illegible]

[illegible]jsonb[illegible]json[illegible]

3.8.4 jsonb[illegible]json[illegible]

[illegible]json[illegible]“->>”[illegible]

```
mydb=> SELECT  name ->> 'col2' FROM test_json1 WHERE id=1;
    ?column?
----------
    francs
(1 row)
```

[illegible]

```
mydb=> SELECT '{"a":1, "b":2}'::jsonb ? 'a';
    ?column?
----------
    t
(1 row)
```

□□json□□□□/□□□□□□□

```
mydb=> SELECT '{"a":1, "b":2}'::jsonb - 'a';
    ?column?
----------
    {"b": 2}
(1 row)
```

3.8.5 jsonb□json□□

json□jsonb□□□□□□□□□□□□□□□□□

□□□□□□json□□□□□□□/□□□□□□□□□□

```
mydb=> SELECT * FROM json_each('{"a":"foo", "b":"bar"}');
    key | value
--------+-------
    a   | "foo"
    b   | "bar"
(2 rows)
```

□□□□□□□□□□□□□□□

```
mydb=> SELECT * FROM json_each_text('{"a":"foo",
"b":"bar"}');
    key | value
--------+-------
    a   | foo
    b   | bar
(2 rows)
```

□□□□□□□□□□row_to_json□□□□□□□□□□□json□□□□□□□□□□□□□json□□□□□□□□□□□□□□□□json□□□□□□□□□□□

```
mydb=> SELECT * FROM test_copy WHERE id=1;
    id | name
-------+------
     1 | a
(1 row)

mydb=> SELECT row_to_json(test_copy) FROM test_copy WHERE
id=1;
    row_to_json
---------------------
    {"id":1,"name":"a"}
(1 row)
```

□□□□□□json□□□□□□□□□□□□□□

```
mydb=> SELECT * FROM json_object_keys('{"a":"foo",
"b":"bar"}');
    json_object_keys
------------------
    a
    b
(2 rows)
```

3.8.6 jsonb□/□□□□□□□□□□

jsonb□/□□□□□□“||”□□□□□□□□sex□/□□□□□□□

```
mydb=> SELECT '{"name":"francs","age":"31"}'::jsonb ||
'{"sex":"male"}'::jsonb;
                    ?column?
------------------------------------------------
    {"age": "31", "sex": "male", "name": "francs"}
(1 row)
```

jsonb□/□□□□□□□□□□□□□□□□□□“-”□□□□□□□□□□□“#-”□□□□□/□□□□□□□“-”□□□/□□□□□□□□□

```
mydb=> SELECT '{"name": "James", "email":
"james@localhost"}'::jsonb
    - 'email';
    ?column?
-------------------
    {"name": "James"}
(1 row)

mydb=> SELECT '["red","green","blue"]'::jsonb - 0;
    ?column?
-------------------
    ["green", "blue"]
```

□□□□□□□□□□□“#-”□□□□□/□□□□□□□□□json□□□□□□□□□□□□□□□□contact□□fax□/□□

```
mydb=> SELECT '{"name": "James", "contact": {"phone": "01234
567890", "fax": "01987 543210"}}'::jsonb #-
'{contact,fax}'::text[];
                        ?column?
------------------------------------------------------------------
    {"name": "James", "contact": {"phone": "01234 567890"}}
(1 row)
```

□□□□aliases□□□□□1□□/□□□□□□□

```
mydb=> SELECT '{"name": "James", "aliases": ["Jamie","The
Jamester","J Man"]}'::jsonb #- '{aliases,1}'::text[];
                     ?column?
--------------------------------------------------
    {"name": "James", "aliases": ["Jamie", "J Man"]}
(1 row)
```

□/□□□□□□□□□□□□□□□□□“||”□□□□“||”□□□□□□□json□□□□□□□□□□□□□□□□□□age□□□□

```
mydb=> SELECT '{"name":"francs","age":"31"}'::jsonb ||
'{"age":"32"}'::jsonb;
            ?column?
---------------------------------
    {"age": "32", "name": "francs"}
(1 row)
```

□□□□□□□□jsonb_set□□□□□□□□

```
jsonb_set(target jsonb, path text[], new_value jsonb[,
create_missing boolean])
```

target□□jsonb□□□path□□□□new_value□□□□□□□□create_missing□□true□□□□□□□□□□□□create_missing□□false□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT
jsonb_set('{"name":"francs","age":"31"}'::jsonb,'{age}','"32
"'::jsonb,false);
            jsonb_set
---------------------------------
    {"age": "32", "name": "francs"}
(1 row)

mydb=> SELECT
jsonb_set('{"name":"francs","age":"31"}'::jsonb,'{sex}','"ma
le"'::jsonb,true);
                   jsonb_set
------------------------------------------------
    {"age": "31", "sex": "male", "name": "francs"}
(1 row)
```

3.9 □□□□□□

□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□CAST□□□□□□□□□□□□□□□□

3.9.1 □□□□□□□□□□□

PostgreSQL□□□□□□□□□□□□□□□□□□3-8□□□

□3-8 □□□□□□□□

函　　数	返回类型	描　　述	示　　例
to_char(timestamp, text)	text	把时间戳转换成字符串	to_char(current_timestamp, 'HH12:MI:SS')
to_char(interval, text)	text	把间隔转换成字符串	to_char(interval '15h 2m 12s', 'HH24:MI:SS')
to_char(int, text)	text	把整数转换成字符串	to_char(125, '999')
to_char(numeric, text)	text	把数字转换成字符串	to_char(-125.8, '999D99S')
to_date(text, text)	date	把字符串转换成日期	to_date('05 Dec 2000', 'DD Mon YYYY')
to_number(text, text)	numeric	把字符串转换成数字	to_number('12,454.8-', '99G999D9S')
to_timestamp(text, text)	timestamp with time zone	把字符串转换成时间戳	to_timestamp('05 Dec 2000', 'DD Mon YYYY')

3.9.2 □□CAST□□□□□□

□varchar□□□□□□□text□□□□□□□□

```
mydb=> SELECT CAST(varchar'123' as text);
    text
------
    123
(1 row)
```

□varchar□□□□□□□int4□□□□□□□□

```
mydb=> SELECT CAST(varchar'123' as int4);
    int4
------
    123
```

3.9.3 □□□□□□□□□□□

□□□□□□□int4□numeric□□□□□□□□

```
mydb=> SELECT 1::int4, 3/2::numeric;
 int4 |      ?column?
------+--------------------
    1 | 1.5000000000000000
(1 row)
```

□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□pg_class□□□□OID□□□OID□□□□□□□□□

```
mydb=> SELECT oid,relname FROM pg_class WHERE
relname='test_json1';
     oid  |  relname
---------+-----------
```

```
    16509 | test_json1
(1 row)
```

[illegible]test_json1[illegible]OID[illegible]pg_attribute[illegible]attrelid[illegible]OID[illegible]

```
mydb=> SELECT attname FROM pg_attribute WHERE
attrelid='16509' AND attnum >0;
    attname
---------
    id
    name
(2 rows)
```

[illegible]

```
mydb=> SELECT attname
        FROM pg_attribute
        WHERE attrelid='test_json1'::regclass AND attnum >0;
    attname
---------
    id
    name
(2 rows)
```

[illegible]

[illegible] pg_class[illegible]PostgreSQL[illegible]OID[illegible]pg_class

[illegible]pg_class[illegible]ID[illegible]pg_attribute [illegible][illegible]pg_attribute.attrelid[illegible]OID[illegible] pgclass.oid[illegible]

3.10 [illegible]

[illegible]PostgreSQL[illegible]/[illegible]json/jsonb[illegible]PostgreSQL[illegible]XML[illegible]https://www.postgresql.org/docs/10/static/datatype.html [illegible]PostgreSQL[illegible]DBA[illegible]

□4□ SQL□□□□

□□□□□PostgreSQL□SQL□□□□□□□□□□WITH□□□□□□□□RET-URNING□□□□□□□□UPSERT□□□□□□□□□□□□□□□□

4.1 WITH□□

WITH□□□PostgreSQL□□□□□SQL□□□□□□□□□□□□CTE□Common Table Expressions□□WITH□□□

4.1.1 □□□□□□CTE

□□□□□□□□CTE□□□□WITH□□□□□□□□

```
WITH t as (
    SELECT generate_series(1,3)
)
SELECT * FROM t;
```

□□□□□□□

```
generate_series
-----------------
              1
              2
              3
(3 rows)
```

□□□□□CTE□□□□□□□□□□□□□□□□t□□□□□□□□□□□□□□t□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□CTE□□□□□□□□□□□□□SQL□

CTE□□□□SQL□□CTE□□□□□□□□□□□□□□□□□

```
WITH regional_sales AS (
        SELECT region, SUM(amount) AS total_sales
        FROM orders
        GROUP BY region
    ), top_regions AS (
        SELECT region
        FROM regional_sales
        WHERE total_sales > (SELECT SUM(total_sales)/10 FROM
regional_sales)
    )
SELECT region,
        product,
        SUM(quantity) AS product_units,
        SUM(amount) AS product_sales
FROM orders
WHERE region IN (SELECT region FROM top_regions)
GROUP BY region, product;
```

□□□□□□□□□regional_sales□top_regions□□□□□□□regional_sales□□□□□□□□□□□□top_regions□□□□□□□□□□10%□□□□□□□□□□□□□□□□□□□□orders□□□□□□□□□□□□□□□□□□□□□□□□

4.1.2 □□□□□□CTE

WITH[illegible]RECURSIVE[illegible]RECURSIVE[illegible]RECURSIVE[illegible]

```
WITH recursive t (x) as (
    SELECT 1
    UNION
    SELECT x + 1
    FROM t
    WHERE x < 5
)
SELECT sum(x) FROM t;
```

[illegible]

```
sum
-----
    15
(1 row)
```

[illegible]x[illegible]1[illegible]union[illegible]1[illegible]x[illegible]5[illegible]x[illegible]

[illegible]PostgreSQL[illegible]

```
id name fatherid
1 [illegible] 0
2 [illegible] 1
3 [illegible] 1
4 [illegible] 2
```

```
5 □□ 2
6 □□ 3
7 □□□ 4
8 □□□ 4
```

□□□□□id□□□□□□□□□□□□□□id=7□□□□□□□□□□□□□□□□□id=5□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□WITH□□□□□□□□□□□□□□□□□□□□□□□

```
CREATE TABLE test_area(id int4,name varchar(32),fatherid
int4);

INSERT INTO test_area VALUES (1, '□□'   ,0);
INSERT INTO test_area VALUES (2, '□□'   ,1);
INSERT INTO test_area VALUES (3, '□□'   ,1);
INSERT INTO test_area VALUES (4, '□□'   ,2);
INSERT INTO test_area VALUES (5, '□□'   ,2);
INSERT INTO test_area VALUES (6, '□□'   ,3);
INSERT INTO test_area VALUES (7, '□□□' ,4);
INSERT INTO test_area VALUES (8, '□□□' ,4);
```

□□PostgreSQL□WITH□□□□ID□7□□□□□□□□□□□□□□□□

```
WITH RECURSIVE r AS (
        SELECT * FROM test_area WHERE id = 7
    UNION   ALL
        SELECT test_area.* FROM test_area, r WHERE
test_area.id = r.fatherid
    )
SELECT * FROM r ORDER BY id;
```

□□□□□□□

```
id |  name  | fatherid
------+-------+----------
    1 | □□   |        0
    2 | □□   |        1
    4 | □□   |        2
    7 | □□□ |       4
(4 rows)
```

□□□□□□□ID=7□□□□□□□□□□□□□□□□□□□□□name□□□□□“□□□□□□□□□”□□□□□□□□□□string_agg□□□□□□□□□□

```
mydb=> WITH RECURSIVE r AS (
        SELECT * FROM test_area WHERE id = 7
    UNION   ALL
        SELECT test_area.* FROM test_area, r WHERE
test_area.id = r.fatherid
    )
SELECT string_agg(name,'') FROM ( SELECT name FROM r ORDER
BY id) n;
    string_agg
--------------------
□□□□□□□□□
```

□□□where□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> WITH RECURSIVE r AS (
        SELECT * FROM test_area WHERE id = 4
```

```
     UNION  ALL
         SELECT test_area.* FROM test_area, r WHERE
 test_area.fatherid = r.id
     )
  SELECT * FROM r ORDER BY id;
  id |  name  | fatherid
----+--------+----------
   4 | □□    |        2
   7 | □□□  |       4
   8 | □□□  |       4
(3 rows)
```

□□□□□CTE□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□CTE□□□□□□□

·CTE□□□□SQL□□□□□SQL□□□□□□□SQL□□□□□□□

·CTE□□□□□□□□□□□□□□□□□□□□□□□□□

·□□□□□□□□□□□□□□□□□□□□

4.2 [illegible]

[illegible]PostgreSQL[illegible]

4.2.1 [illegible]INSERT INTO...SELECT...

[illegible]

```
INSERT INTO table_name SELECT...FROM source_table
```

[illegible]user_ini[illegible]user_ini[illegible]

```
mydb=> CREATE TABLE tbl_batch1(user_id int8,user_name text);
CREATE TABLE

mydb=> INSERT INTO tbl_batch1(user_id,user_name)
SELECT user_id,user_name FROM user_ini;
INSERT 0 1000000
```

[illegible]user_ini[illegible]user_id[illegible]user_name[illegible]tbl_batch1[illegible]where[illegible]

[illegible]

```
mydb=> CREATE TABLE tbl_batch2 (id int4,info text);
CREATE TABLE

mydb=> INSERT INTO tbl_batch2(id,info)
SELECT generate_series(1,5),'batch2';
INSERT 0 5
```

[illegible]SELECT[illegible]PostgreSQL[illegible]

4.2.2 [illegible]INSERT INTO VALUES[illegible]…[illegible]

PostgreSQL[illegible]INSERT[illegible]VALUES[illegible]

```
mydb=> CREATE TABLE tbl_batch3(id int4,info text);
CREATE TABLE

mydb=> INSERT INTO tbl_batch3(id,info) VALUES (1,'a'),
(2,'b'),(3,'c');
INSERT 0 3
```

[illegible]

```
mydb=> SELECT * FROM tbl_batch3;
    id | info
-------+------
     1 | a
     2 | b
```

```
      3 | c
(3 rows)
```

□□□□□□□□□□□□□□□SQL□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□WAL□Write-Ahead Logging□□□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□

4.2.3 □□□□COPY□\COPY□□□

2.2.3□□□□psql□□□□□□□□□□□□□COPY□□□\copy□□□□copy□\copy□□□□□□□□□□□□□□□□□□□□□□□□□INSERT□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□□COPY□□□□□□□□□□□□□□□□□COPY□□□□□□□□□□□□□□□□□□□□□4□CPU□8GB□□□

□□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> CREATE TABLE tbl_batch4(
id int4,
info text,
create_time timestamp(6) with time zone default
clock_timestamp());
CREATE TABLE

mydb=> INSERT INTO tbl_batch4(id,info) SELECT n,n||'_batch4'
FROM generate_series(1,10000000) n;
INSERT 0 10000000
```

[illegible]INSERT[illegible]

```
[postgres@pghost1 ~]$ psql mydb postgres
psql (10.0)
Type "help" for help.

mydb=# \timing
Timing is on.

mydb=# COPY pguser.tbl_batch4 TO '/home/pg10/tbl_batch4.txt';
COPY 10000000
Time: 6575.787 ms (00:06.576)
```

[illegible]6575[illegible]tbl_batch4[illegible]tbl_batch4.txt[illegible]

```
mydb=# TRUNCATE TABLE pguser.tbl_batch4;
TRUNCATE TABLE
mydb=# COPY pguser.tbl_batch4 FROM
'/home/pg10/tbl_batch4.txt';
COPY 10000000
Time: 15663.834 ms (00:15.664)
```

[illegible]COPY[illegible]15663[illegible]

4.3 RETURNING□□□□□□□

PostgreSQL□RETURNING□□□□□□DML□□□□□□□□□□□□□□INSERT□□□□RETURNING□□□□□□□□□□□UPDATE□□□□RETURNING□□□□□□□□□□□DELETE□□□□RETURNING□□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□

4.3.1 RETURNING□□□□□□□

INSERT□□□□RETURNING□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> CREATE TABLE test_r1(id serial,flag char(1));
CREATE TABLE

mydb=> INSERT INTO test_r1(flag) VALUES ('a') RETURNING *;
    id | flag
-------+------
     1 | a
(1 row)
INSERT 0 1
```

RETURNING*□□□□□□□□□□□□□□□□□□□□□□□□□RETURNING□□□□□□□□□□□□□□□□□□id□□□

```
mydb=> INSERT INTO test_r1(flag) VALUES ('b') RETURNING id;
    id
----
     2
(1 row)
INSERT 0 1
```

4.3.2 RETURNING□□□□□□□

UPDATE□□RETURNING□□□□UPDATE□□□□□□□□□□□□□

```
mydb=> SELECT * FROM test_r1 WHERE id=1;
    id | flag
-------+------
     1 | a
(1 row)

mydb=> UPDATE test_r1 SET flag='p' WHERE id=1 RETURNING *;
    id | flag
-------+------
     1 | p
(1 row)
UPDATE 1
```

4.3.3 RETURNING□□□□□□□

DELETE□□RETURNING□□□□□□□□□□□□□□□

```
mydb=> DELETE FROM test_r1 WHERE id=2 RETURNING *;
    id | flag
-------+------
     2 | b
(1 row)
DELETE 1
```

4.4 UPSERT

PostgreSQL□UPSERT□□□□INSERT...ON CONFLICT UPDATE□□PostgreSQL□UPSERT□□□□□□□□□□

4.4.1 UPSERT□□□□

□□□□□□□□□□UPSERT□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> CREATE TABLE user_logins(user_name text primary key,
login_cnt int4,
last_login_time timestamp(0) without time zone);
CREATE TABLE

mydb=> INSERT INTO user_logins(user_name,login_cnt) VALUES
('francs',1);
INSERT 0 1
```

□user_logins□user_name□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> INSERT INTO user_logins(user_name,login_cnt)
VALUES ('matiler',1),('francs',1);
```

```
ERROR:  duplicate key value violates unique constraint
"user_logins_pkey"
DETAIL:  Key (user_name)=(francs) already exists.
```

□□SQL□□□□□□□□□□□matiler□□□□□□□□□□□□francs□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□UPSERT□□

```
mydb=> INSERT INTO user_logins(user_name,login_cnt)
VALUES ('matiler',1),('francs',1)
ON CONFLICT(user_name)
DO UPDATE SET
login_cnt=user_logins.login_cnt+EXCLUDED.login_cnt,last_logi
n_time=now();
INSERT 0 2
```

□□INSERT□□□□□□□□□□□□□□□□□□□□□□□□□□□□login_cnt□□1□□□□□□□□□□last_login_time□ON CONFLICT □user_name□□□□□□□□user_name□□□DO UPDATE SET□□□□□□□□□□□□□□UPDATE□□□□□□□SET□□□□□□user_loins□□□□□EXCLUDED□□□□□user_loins□□□□□□□□□□□□□□□□EXCLUDED□□□□□□□□□□□□□□user_login□□□□□□

```
mydb=> SELECT * FROM user_logins ;
    user_name | login_cnt |   last_login_time
```

```
-------------+----------+--------------------
    matiler   |        1 |
    francs    |        2 | 2017-08-08 15:23:13
(2 rows)
```

□□□□□□francs□□□□□□□□login_cnt□last_login_time□□□□□□□□□□□matiler□□□□□□□□

□□□□□□□□□□□□□□□□□□□□DO NOTHING□□□□□□□□

```
mydb=> INSERT INTO user_logins(user_name,login_cnt)
VALUES ('tutu',1),('francs',1)
ON CONFLICT(user_name) DO NOTHING;
INSERT 0 1
```

□□□□□□□□□□□□tutu□□□□□□□□□□□□□□francs□□□□□□□□□□□□□□

```
mydb=> SELECT * FROM user_logins ;
    user_name | login_cnt |   last_login_time
-------------+----------+--------------------
    matiler   |        1 |
    francs    |        2 | 2017-08-08 15:23:13
    tutu      |        1 |
(3 rows)
```

4.4.2 UPSERT□□

PostgreSQL□UPSERT□□□□□□□□□□□□□□
□□□□□□□□□□□□□□

```
INSERT INTO table_name [ AS alias ] [ ( column_name [, ...] ) ]
    [ ON CONFLICT [ conflict_target ] conflict_action ]

where conflict_target can be one of:
    ( { index_column_name | ( index_expression ) } [ COLLATE collation ] [ opclass ] [, ...] ) [ WHERE index_predicate ]
    ON CONSTRAINT constraint_name

and conflict_action is one of:
    DO NOTHING
    DO UPDATE SET { column_name = { expression | DEFAULT } |
                    ( column_name [, ...] ) = [ ROW ] ( { expression | DEFAULT } [, ...] ) |
                    ( column_name [, ...] ) = ( sub-SELECT )
                  } [, ...]
          [ WHERE condition ]
```

□□□□□□□□[ON CONFLICT[conflict_target]conflict_action]□□□conflict_target□□□□□□□□□□□□□□□□□□□□□□□□□□□□conflict_action□□□□□□□□□DO NOTHING□□□□□□□□□□□UPDATE□□□

4.5 [illegible]

[illegible]TABLESAMPLE[illegible]PostgreSQL[illegible]9.5[illegible]TABLESAMPLE[illegible]9.5[illegible]ORDER BY random[illegible]

```
mydb=> SELECT * FROM user_ini ORDER BY random() LIMIT 1;
        id   | user_id | user_name |          create_time
-------------+---------+-----------+-------------------------------
    500449   |  768810 | 2TY6P4    | 2017-08-05
15:59:32.294761+08
(1 row)

mydb=> SELECT * FROM user_ini ORDER BY random() LIMIT 1;
        id   | user_id | user_name |          create_time
-------------+---------+-----------+-------------------------------
      324823 |  740720 | 07SKCU    | 2017-08-05
15:59:29.913984+08
(1 row)
```

[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT * FROM user_ini ORDER BY
random() LIMIT 1;
                                    QUERY PLAN
--------------------------------------------------------------------------------
```

```
 Limit  (cost=25599.98..25599.98 rows=1 width=35) (actual
time=367.867..367.868 rows=1 loops=1)
    ->  Sort  (cost=25599.98..28175.12 rows=1030056
width=35) (actual time= 367.866..367.866 rows=1 loops=1)
        Sort Key: (random())
        Sort Method: top-N heapsort  Memory: 25kB
            ->  Seq Scan on user_ini  (cost=0.00..20449.70
rows=1030056 width=35) (actual time=0.012..159.569
rows=1000000 loops=1)
 Planning time: 0.083 ms
 Execution time: 367.909 ms
(7 rows)
```

□user_ini□□□□100□□□100□□□□□□□□SQL□□□□□□367ms□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

9.5□□□□PostgreSQL□□TABLESAMPLE□□□□□□□□□□□□

```
SELECT ...
FROM table_name
TABLESAMPLE sampling_method ( argument [, ...] ) [
REPEATABLE ( seed ) ]
```

sampling_method□□□□□□□□□□□□SYSTEM□BERNOULLI□□□□□□□□□□□□□□□□argument□□□□□□□

□□ explain analyze□□□□□□□□□□SQL□□□□□SQL□□□□□□□□□□Planning time□□

SQL[illegible]Execution time[illegible]SQL[illegible]

4.5.1 SYSTEM[illegible]

SYSTEM[illegible]SYSTEM[illegible]

[illegible]test_sample[illegible]150[illegible]

```
mydb=> CREATE TABLE test_sample(id int4,message text,
create_time timestamp(6) without time zone default
clock_timestamp());
CREATE TABLE

mydb=> INSERT INTO test_sample(id,message)
SELECT n, md5(random()::text) FROM
generate_series(1,1500000) n;
INSERT 0 1500000

mydb=>  SELECT * FROM test_sample LIMIT 1;
    id |             message              |
create_time
-------+----------------------------------+-----------------
-----------
     1 | 58f2506410be948963d6d9adf4b4e0c2 | 2017-08-08
21:17:20.984481
(1 row)
```

[illegible]0.01[illegible]1500000×0.01%=150[illegible]SQL[illegible]

```
EXPLAIN ANALYZE SELECT * FROM test_sample TABLESAMPLE
SYSTEM(0.01);
```

[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT * FROM test_sample TABLESAMPLE
SYSTEM(0.01);
                                  QUERY PLAN
-----------------------------------------------------------------------------
    Sample Scan on test_sample  (cost=0.00..3.50 rows=150
width=45) (actual time=0.099..0.146 rows=107 loops=1)
        Sampling: system ('0.01'::real)
    Planning time: 0.053 ms
    Execution time: 0.166 ms
(4 rows)
```

[illegible]Sample Scan[illegible]SYSTEM[illegible]0.166[illegible]150[illegible]107[illegible]107[illegible]

```
mydb=> SELECT relname,relpages FROM pg_class WHERE
relname='test_sample';
        relname | relpages
----------------+----------
    test_sample |    14019
(1 row)
```

[illegible]test_sample[illegible]14019[illegible]1000000/14019=107[illegible]

[illegible]ctid[illegible]

```
mydb=> SELECT ctid,* FROM test_sample TABLESAMPLE
SYSTEM(0.01);
    ctid    |   id   |             message              |
create_time
------------+--------+----------------------------------+---
-----------------------
 (5640,1)   | 603481 | 385484b3452b245e46388d71ce4ea928 |
2017-08-08 21:17:23.32394
 (5640,2)   | 603482 | e09c526118f1d4b3c391d59ae915c4e8 |
2017-08-08 21:17:23.323964
….[illegible]
(5640,107) | 603587 | c33875a052f4ca63c4b38c649fb6bcc3 |
2017-08-08 21:17:23.324336
(107 rows)
```

ctid[illegible]107[illegible]5640[illegible]0.01[illegible]

```
mydb=>  SELECT count(*) FROM test_sample TABLESAMPLE
SYSTEM(0.01);
    count
-------
    214
(1 row)

mydb=>  SELECT count(*) FROM test_sample TABLESAMPLE
SYSTEM(0.01);
    count
```

```
-------
    107
(1 row)
```

[illegible]214[illegible]107[illegible]107[illegible]0.01[illegible]1500000×0.01%=150[illegible]150[illegible]SYSTEM[illegible]

4.5.2 BERNOULLI[illegible]

BERNOULLI[illegible]BERNOULLI[illegible]BERNOULLI[illegible]SYSTEM[illegible]SYSTEM[illegible]BERNOULLI[illegible]test_sample[illegible]

[illegible]BERNOULLI[illegible]0.01[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT * FROM test_sample TABLESAMPLE
BERNOULLI (0.01);
                                     QUERY PLAN
-------------------------------------------------------------
--------------------
    Sample Scan on test_sample  (cost=0.00..14020.50
rows=150 width=45) (actual time=0.025..22.541 rows=152
loops=1)
        Sampling: bernoulli ('0.01'::real)
```

```
    Planning time: 0.063 ms
    Execution time: 22.569 ms
(4 rows)
```

[illegible]Sample Scan[illegible]BERNOULLI[illegible]150[illegible]152[illegible]150[illegible]1000000×0.01%[illegible]22.569[illegible]SYSTEM[illegible]0.166[illegible]136[illegible]

[illegible]

```
mydb=>  SELECT count(*) FROM test_sample TABLESAMPLE
BERNOULLI(0.01);
    count
-------
    151
(1 row)

mydb=>  SELECT count(*) FROM test_sample TABLESAMPLE
BERNOULLI(0.01);
    count
-------
    147
(1 row)
```

[illegible]BERNOULLI[illegible]SYSTEM[illegible]SYSTEM[illegible]

□□BERNOULLI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ctid□□□□□□□□□□

```
mydb=>  SELECT ctid,id,message
        FROM test_sample TABLESAMPLE BERNOULLI(0.01) lIMIT
3;
    ctid   |  id   |             message
-----------+-------+----------------------------------
  (55,30)  |  5915 | f3803f234f6cf6cdd276d9d027487582
  (240,23) | 25703 | c04af69ac76f6465832e0cd87939a1af
  (318,3)  | 34029 | dd35438b24980d1a8ed2d3f5edd5ca1c
```

□□□□□□□□ctid□□□□□□□□□□□□□□□□55□240□318□□□□BERNOULLI□□□□□□□□□SYSTEM□□□□□□□

□□□□□SYSTEM□BERNOULLI□□□□□SYSTEM□□GB□□□□□□BERNOULLI□□□□□□□□□□□□SYSTEM□□□□□□□□□□□□□□□□□□□□SYSTEM□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

4.6 □□□□

□□□□□□□□□□□□□□□□□□□□□□□avg□□□sum□□□min□□□max□□□count□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□ □□
□□ □□
□□ □□
□□ □□
□□ □□

□□□□□□□□□□

□□ □□□□□□□□
□□ □□□□□

□□□□□□SQL□□□□

4.6.1 string_agg□□

□□□□string_agg□□□□□□□□□□□□□

```
string_agg(expression, delimiter)
```

□□□□string_agg□□□□□□□□□□□□□□□□□□□□□□□□□□□□delimiter□□□□□□expression□□□□□□□□□□□□□□□□□□□□text□text□□□bytea□bytea□□□□□□□□□□□□□□□□□□□bytea□□□□□□□□□□□□□□□□□□□□□□text□□□□□□□□□□□□□□□□□□□string_agg□□□□□

□□□□□□□□□□□□□□□

```
CREATE TABLE city (country character varying(64),city
character varying(64));
INSERT INTO city VALUES ('□□','□□');
INSERT INTO city VALUES ('□□','□□');
INSERT INTO city VALUES ('□□','□□');
INSERT INTO city VALUES ('□□','□□');
INSERT INTO city VALUES ('□□','□□');
```

□□□□□□□

```
mydb=> SELECT * FROM city;
    country | city
------------+------
    □□    | □□
    □□    | □□
    □□    | □□
    □□    | □□
    □□    | □□
(5 rows)
```

□city□□□□□□□□□□□□□□□□

```
mydb=> SELECT string_agg(city,',') FROM city;
        string_agg
--------------------------
    □□,□□,□□,□□,□□
(1 row)
```

□□string_agg□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□

```
mydb=> SELECT country,string_agg(city,',') FROM city GROUP
BY country;
    country |   string_agg
------------+---------------
    □□      | □□,□□
    □□      | □□,□□,□□
(2 rows)
```

4.6.2　array_agg□□

array_agg□□□string_agg□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□array_agg
□□□□□□□□□□□□□□□□□

```
array_agg(expression)  --□□□□□□□□□□□□
```

[illegible]array_agg[illegible]

```
mydb=> SELECT country,array_agg(city) FROM city GROUP BY
country;
    country |    array_agg
------------+-----------------
    [illegible]     | {[illegible],[illegible]}
    [illegible]     | {[illegible],[illegible],[illegible]}
```

array_agg[illegible]array_agg[illegible]

[illegible]array_agg[illegible]

```
array_agg(expression)  --[illegible]
```

[illegible]array_agg[illegible]

[illegible]

```
mydb=> CREATE TABLE test_array3(id int4[]);
CREATE TABLE
mydb=> INSERT INTO test_array3(id) VALUES (array[1,2,3]);
INSERT 0 1
mydb=> INSERT INTO test_array3(id) VALUES (array[4,5,6]);
INSERT 0 1
```

[illegible]

```
mydb=> SELECT * FROM test_array3;
        id
---------
    {1,2,3}
    {4,5,6}
(2 rows)
```

[illegible]array_agg[illegible]

```
mydb=> SELECT  array_agg(id) FROM test_array3;
        array_agg
-------------------
    {{1,2,3},{4,5,6}}
(1 row)
```

[illegible]array_agg[illegible]array_to_string[illegible]

```
mydb=> SELECT array_to_string( array_agg(id),',') FROM
test_array3;
    array_to_string
-----------------
    1,2,3,4,5,6
(1 row)
```

4.7 [illegible]

[illegible]SQL[illegible]

4.7.1 [illegible]

PostgreSQL[illegible]row_num[illegible]rank[illegible]lag[illegible]OVER[illegible]

[illegible]

```
function_name ([expression [, expression ... ]]) [ FILTER (
WHERE filter_clause ) ] OVER ( window_definition )
```

[illegible]window_definition[illegible]

```
[ existing_window_name ]
[ PARTITION BY expression [, ...] ]
[ ORDER BY expression [ ASC | DESC | USING operator ] [
NULLS { FIRST | LAST } ] [, ...] ]
[ frame_clause ]
```

[illegible]

·OVER[illegible]

·PARTITON BY[illegible]

·ORDER BY[illegible]

[illegible]

4.7.2 avg[illegible]OVER[illegible]

[illegible]OVER[illegible]avg[illegible]OVER[illegible]

[illegible]

```
CREATE TABLE score ( id serial primary key,
                      subject character varying(32),
                      stu_name character varying(32),
                      score numeric(3,0) );

INSERT INTO score ( subject,stu_name,score ) VALUES
('Chinese','francs',70);
INSERT INTO score ( subject,stu_name,score ) VALUES
('Chinese','matiler',70);
INSERT INTO score ( subject,stu_name,score) VALUES
('Chinese','tutu',80);
INSERT INTO score ( subject,stu_name,score ) VALUES
```

```
('English','matiler',75);
INSERT INTO score ( subject,stu_name,score ) VALUES
('English','francs',90);
INSERT INTO score ( subject,stu_name,score ) VALUES
('English','tutu',60);
INSERT INTO score ( subject,stu_name,score ) VALUES
('Math','francs',80);
INSERT INTO score ( subject,stu_name,score ) VALUES
('Math','matiler',99);
INSERT INTO score ( subject,stu_name,score ) VALUES
('Math','tutu',65);
```

[illegible]score[illegible]

```
mydb=> SELECT s.subject, s.stu_name,s.score, tmp.avgscore
    FROM score s
    LEFT JOIN (SELECT subject, avg(score) avgscore FROM
score GROUP BY subject) tmp
        ON s.subject = tmp.subject;
    subject | stu_name | score |      avgscore
------------+----------+-------+---------------------
    Chinese | francs   |    70 | 73.3333333333333333
    Chinese | matiler  |    70 | 73.3333333333333333
    Chinese | tutu     |    80 | 73.3333333333333333
    English | matiler  |    75 | 75.0000000000000000
    English | francs   |    90 | 75.0000000000000000
    English | tutu     |    60 | 75.0000000000000000
    Math    | francs   |    80 | 81.3333333333333333
    Math    | matiler  |    99 | 81.3333333333333333
    Math    | tutu     |    65 | 81.3333333333333333
(9 rows)
```

[illegible]

```
mydb=> SELECT subject,stu_name, score, avg(score)
OVER(PARTITION BY subject) FROM score;
```

```
    subject | stu_name | score |         avg
------------+----------+-------+---------------------
    Chinese | francs   |    70 | 73.3333333333333333
    Chinese | matiler  |    70 | 73.3333333333333333
    Chinese | tutu     |    80 | 73.3333333333333333
    English | matiler  |    75 | 75.0000000000000000
    English | francs   |    90 | 75.0000000000000000
    English | tutu     |    60 | 75.0000000000000000
    Math    | francs   |    80 | 81.3333333333333333
    Math    | matiler  |    99 | 81.3333333333333333
    Math    | tutu     |    65 | 81.3333333333333333
(9 rows)
```

□□□□□□□□□□□□score□□□□□□□□□□□□□□□PARTITION BY subject□□□□□□subject□□□□□

4.7.3 row_number□□

row_number□□□□□□□□□□□□□□□□□□□□□□□□1□□□□□□□□□

```
mydb=> SELECT row_number() OVER (partition by subject ORDER
BY score desc),* FROM score;
    row_number| id | subject | stu_name | score
--------------+----+---------+----------+-------
            1 |  3 | Chinese | tutu     |    80
            2 |  1 | Chinese | francs   |    70
            3 |  2 | Chinese | matiler  |    70
            1 |  5 | English | francs   |    90
            2 |  4 | English | matiler  |    75
            3 |  6 | English | tutu     |    60
            1 |  8 | Math    | matiler  |    99
            2 |  7 | Math    | francs   |    80
            3 |  9 | Math    | tutu     |    65
(9 rows)
```

[illegible]row_number[illegible]partition[illegible]row_number[illegible]oracle[illegible]ROWNUM[illegible]

```
mydb=> SELECT  row_number() OVER (ORDER BY id) AS rownum ,*
FROM score;
    rownum | id | subject | stu_name | score
-----------+----+---------+----------+-------
         1 |  1 | Chinese | francs   |    70
         2 |  2 | Chinese | matiler  |    70
         3 |  3 | Chinese | tutu     |    80
         4 |  4 | English | matiler  |    75
         5 |  5 | English | francs   |    90
         6 |  6 | English | tutu     |    60
         7 |  7 | Math    | francs   |    80
         8 |  8 | Math    | matiler  |    99
         9 |  9 | Math    | tutu     |    65
(9 rows)
```

4.7.4 rank[illegible]

rank[illegible]row_number[illegible]gaps[illegible]

```
mydb=> SELECT rank() OVER(PARTITION BY subject ORDER BY
score),* FROM score;
    rank | id | subject | stu_name | score
---------+----+---------+----------+-------
       1 |  2 | Chinese | matiler  |    70
       1 |  1 | Chinese | francs   |    70
       3 |  3 | Chinese | tutu     |    80
       1 |  6 | English | tutu     |    60
       2 |  4 | English | matiler  |    75
```

```
        3 |  5 | English | francs   |    90
        1 |  9 | Math    | tutu     |    65
        2 |  7 | Math    | francs   |    80
        3 |  8 | Math    | matiler  |    99
(9 rows)
```

□□□□□□Chinese□□□□□□□□score□□□□□70□□□□□□□rank□□□□1□□□□□□rank□□□□3□□□□□□□

4.7.5 dense_rank□□

dense_rank□□□□□□□rank□□

```
mydb=> SELECT dense_rank() OVER(PARTITION BY subject ORDER
BY score),* FROM score;
    dense_rank | id | subject | stu_name | score
---------------+----+---------+----------+-------
             1 |  2 | Chinese | matiler  |    70
             1 |  1 | Chinese | francs   |    70
             2 |  3 | Chinese | tutu     |    80
             1 |  6 | English | tutu     |    60
             2 |  4 | English | matiler  |    75
             3 |  5 | English | francs   |    90
             1 |  9 | Math    | tutu     |    65
             2 |  7 | Math    | francs   |    80
             3 |  8 | Math    | matiler  |    99
(9 rows)
```

□□□□□□Chinese□□□□□□rank□□□1□□□□□□rank□□□□2□□□□□□□□

4.7.6 lag□□

□□□□□□□□□lag□□□□□□□□□□offset□□□□□□□□□□□□□□□

```
lag(value anyelement [, offset integer [, default anyelement
]])
```

□□□

·value□□□□□□□□□□□

·offset□□1□

·default□□□□□□□offset□□□□□□□□□□□□default□□□□null□

□□□□□score□□□□□□□□□□□□□id□□□□□□□

```
mydb=> SELECT lag(id,1)OVER(),* FROM score;
    lag | id | subject | stu_name | score
--------+----+---------+----------+-------
        |  1 | Chinese | francs   |    70
      1 |  2 | Chinese | matiler  |    70
      2 |  3 | Chinese | tutu     |    80
      3 |  4 | English | matiler  |    75
      4 |  5 | English | francs   |    90
      5 |  6 | English | tutu     |    60
      6 |  7 | Math    | francs   |    80
```

```
       7 |  8 | Math     | matiler  |    99
       8 |  9 | Math     | tutu     |    65
(9 rows)
```

□□score□□□□□□□□□□□□□id□□□□□□□□□□□□□□□□

```
mydb=> SELECT lag(id,2,1000)OVER(),* FROM score;
    lag  | id | subject | stu_name | score
---------+----+---------+----------+-------
    1000 |  1 | Chinese | francs   |    70
    1000 |  2 | Chinese | matiler  |    70
       1 |  3 | Chinese | tutu     |    80
       2 |  4 | English | matiler  |    75
       3 |  5 | English | francs   |    90
       4 |  6 | English | tutu     |    60
       5 |  7 | Math    | francs   |    80
       6 |  8 | Math    | matiler  |    99
       7 |  9 | Math    | tutu     |    65
(9 rows)
```

□□□□□lag□□□□□□□□□□□□□□□□□□□offset□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

4.7.7 first_value□□

first_value□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□score□□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT first_value(score) OVER( PARTITION BY subject
),* FROM score;
    first_value | id | subject | stu_name | score
----------------+----+---------+----------+-------
             70 |  1 | Chinese | francs   |    70
             70 |  2 | Chinese | matiler  |    70
             70 |  3 | Chinese | tutu     |    80
             75 |  4 | English | matiler  |    75
             75 |  5 | English | francs   |    90
             75 |  6 | English | tutu     |    60
             80 |  7 | Math    | francs   |    80
             80 |  8 | Math    | matiler  |    99
             80 |  9 | Math    | tutu     |    65
(9 rows)
```

□□first_value□□□□□□□□□□□□□□□□□□□□□ □□□□□score□□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT first_value(score) OVER( PARTITION BY subject
ORDER BY score desc),* FROM score;
    first_value | id | subject | stu_name | score
----------------+----+---------+----------+-------
             80 |  3 | Chinese | tutu     |    80
             80 |  1 | Chinese | francs   |    70
             80 |  2 | Chinese | matiler  |    70
             90 |  5 | English | francs   |    90
             90 |  4 | English | matiler  |    75
             90 |  6 | English | tutu     |    60
             99 |  8 | Math    | matiler  |    99
             99 |  7 | Math    | francs   |    80
             99 |  9 | Math    | tutu     |    65
(9 rows)
```

4.7.8 last_value□□

last_value□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□

□□score□□□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT last_value(score) OVER( PARTITION BY subject
),* FROM score;
 last_value | id | subject | stu_name | score
------------+----+---------+----------+-------
         80 |  1 | Chinese | francs   |    70
         80 |  2 | Chinese | matiler  |    70
         80 |  3 | Chinese | tutu     |    80
         60 |  4 | English | matiler  |    75
         60 |  5 | English | francs   |    90
         60 |  6 | English | tutu     |    60
         65 |  7 | Math    | francs   |    80
         65 |  8 | Math    | matiler  |    99
         65 |  9 | Math    | tutu     |    65
(9 rows)
```

4.7.9 nth_value□□

nth_value□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□

```
nth_value(value any, nth integer)
```

□□□

value□□□□□□□

nth[illegible]

[illegible]score[illegible]

```
mydb=> SELECT nth_value(score,2) OVER( PARTITION BY subject
),* FROM score;
    nth_value | id | subject | stu_name | score
-------------+----+---------+----------+-------
           70 |  1 | Chinese | francs   |    70
           70 |  2 | Chinese | matiler  |    70
           70 |  3 | Chinese | tutu     |    80
           90 |  4 | English | matiler  |    75
           90 |  5 | English | francs   |    90
           90 |  6 | English | tutu     |    60
           99 |  7 | Math    | francs   |    80
           99 |  8 | Math    | matiler  |    99
           99 |  9 | Math    | tutu     |    65
(9 rows)
```

4.7.10 [illegible]

[illegible]SQL[illegible]

```
SELECT .. FROM .. WINDOW window_name AS ( window_definition
) [, ...]
```

WINDOW[illegible]window_name[illegible]OVER[illegible]

```
mydb=> SELECT avg(score) OVER(r),sum(score) OVER(r),* FROM
SCORE WINDOW r as (PARTITION BY subject);
        avg             | sum | id | subject | stu_name |
score
------------------------+-----+----+---------+----------+---
----
    73.3333333333333333 | 220 |  1 | Chinese | francs   |
70
    73.3333333333333333 | 220 |  2 | Chinese | matiler  |
70
    73.3333333333333333 | 220 |  3 | Chinese | tutu     |
80
    75.0000000000000000 | 225 |  4 | English | matiler  |
75
    75.0000000000000000 | 225 |  5 | English | francs   |
90
    75.0000000000000000 | 225 |  6 | English | tutu     |
60
    81.3333333333333333 | 244 |  7 | Math    | francs   |
80
    81.3333333333333333 | 244 |  8 | Math    | matiler  |
99
    81.3333333333333333 | 244 |  9 | Math    | tutu     |
65
(9 rows)
```

[illegible]

4.8 □□□□

□□□□□PostgreSQL□□□□□□□SQL□□□□□□□□□□□□SQL□□□SQL□□□□□□□□□PostgreSQL□□□□□SQL□□□

[illegible]

[illegible]5[illegible] [illegible]

PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]

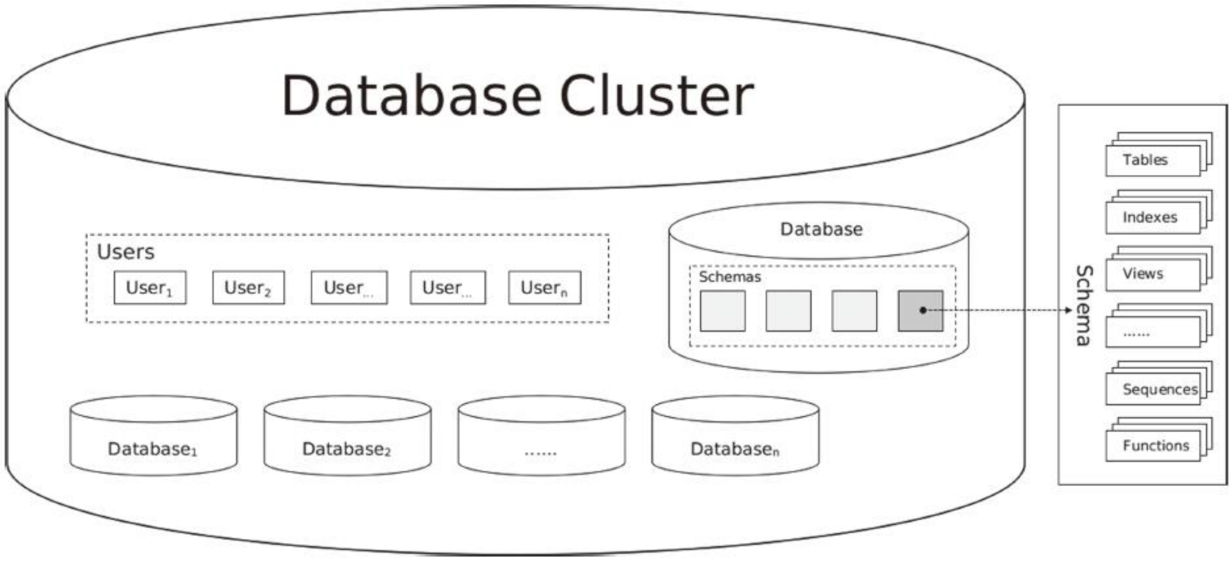
5.1 [illegible]

[illegible]PostgreSQL[illegible]Database Cluster[illegible]PostgreSQL[illegible]“[illegible]”[illegible]PostgreSQL[illegible]

5.1.1 [illegible]

[illegible]PostgreSQL[illegible]

[illegible]Database[illegible]User[illegible]Database[illegible]Database[illegible]User[illegible]5-1[illegible]



□5-1 PostgreSQL□□□□□□□□□

□□□□Database□□□□□Database□□□□□□public□□□Schema□□□Database□□□□□Schema□□□□□□□□□□□□□□□□□□□□□□□□Schema□□□□public□□Schema□□Schema□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Schema□□□□□□□□□□□□□□□□□□□□□□□□□Schema□□□□Schema□□□□□□□□□□□Table□Index□View□Sequence□Function□□□□□□□

5.1.2 □□□□□□

□□□□□□□□□□□initdb□□

1.□□□□□□

□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□PGDATA□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ tree -L 1 -d /pgdata/10/data
/pgdata/10/data
├── base
├── pg_tblspc
├── ...
├── ...
├── ...
├── pg_wal
└── global
```

□5-1□□□□□□□□□□□□□□□□□□□□□

□5-1 □□□□□□□□□□□□□□

目　录	用　途
base	包含每个数据库对应的子目录的子目录
global	包含集簇范围的表的子目录，比如 pg_database
pg_commit_ts	包含事务提交时间戳数据的子目录
pg_xact	包含事务提交状态数据的子目录
pg_dynshmem	包含被动态共享内存子系统所使用文件的子目录
pg_logical	包含用于逻辑复制的状态数据的子目录
pg_multixact	包含多事务状态数据的子目录（用于共享的行锁）
pg_notify	包含 LISTEN/NOTIFY 状态数据的子目录
pg_repslot	包含复制槽数据的子目录
pg_serial	包含已提交的可序列化事务信息的子目录
pg_snapshots	包含导出的快照的子目录
pg_stat	包含用于统计子系统的永久文件的子目录
pg_stat_tmp	包含用于统计信息子系统临时文件的子目录
pg_subtrans	包含子事务状态数据的子目录
pg_tblspc	包含指向表空间的符号链接的子目录
pg_twophase	用于预备事务状态文件的子目录
pg_wal	保存预写日志
pg_xact	记录事务提交状态数据
文　件	**用　途**
PG_VERSION	PostgreSQL 主版本号文件
pg_hba.conf	客户端认证控制文件
postgresql.conf	参数文件
postgresql.auto.conf	参数文件，只保存 ALTER SYSTEM 命令修改的参数
postmaster.opts	记录服务器最后一次启动时使用的命令行参数

2.□□□□□□

□□□□□□base□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□base□□□□□□□□□□□□□□□□□□□□□OID□□□□□

□1□OID

PostgreSQL[illegible]OID[illegible]4[illegible]OID[illegible]OID[illegible]pg_database[illegible]OID[illegible]

```
SELECT oid,datname FROM pg_database WHERE datname = 'mydb';
    oid   | datname
    ------+---------
    16384 | mydb
    (1 row)
```

[illegible]OID[illegible]pg_class[illegible]OID[illegible]

```
mydb=# SELECT oid,relname,relkind FROM pg_class WHERE
relname ~ 'tbl';
    oid   |              relname              | relkind
----------+-----------------------------------+---------
    16385 | tbl_id_seq                        | S
    16387 | tbl                               | r
    16396 | tbl_pkey                          | i
     3455 | pg_class_tblspc_relfilenode_index | i
(4 rows)
```

[illegible]2[illegible]

[illegible]PostgreSQL[illegible]

[illegible]
pg_default[illegible]pg_global[illegible]

```
mydb=# \db
            List of tablespaces
        Name    |  Owner   | Location
---------------+----------+----------
    pg_default | postgres |
    pg_global  | postgres |
(2 rows)
```

·pg_global[illegible]global[illegible]

·pg_default[illegible]base[illegible]template0[illegible]template1[illegible]template1[illegible]template1[illegible]
pg_default[illegible]

[illegible]

·[illegible]

·[illegible]WAL[illegible]

[illegible]postgres[illegible]CREATE TABLESPACE[illegible]

```
[postgres@pghost1 ~]$ mkdir -p /pgdata/10/mytblspc
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/psql -p 1921 mydb
psql (10.2)
Type "help" for help.
mydb=# CREATE TABLESPACE myspc LOCATION
'/pgdata/10/mytblspc';
CREATE TABLESPACE
mydb=# \db
            List of tablespaces
        Name   |  Owner   |      Location
--------------+----------+--------------------
    myspc      | postgres | /pgdata/10/mytblspc
    pg_default | postgres |
    pg_global  | postgres |
(3 rows)
```

[illegible]

```
mydb=# CREATE TABLE t(id SERIAL PRIMARY KEY, ival int)
TABLESPACE myspc;
CREATE TABLE
```

[illegible]OID[illegible]mydb[illegible]pg_database[illegible]OID[illegible]

```
mydb=# SELECT oid,datname FROM pg_database WHERE datname =
'mydb';
      oid | datname
---------+---------
    16384 | mydb
(1 row)
```

□□□□□□□mydb□OID□16384□□□□□□□mydb□□□□□□□□□$PGDATA/base/16384□□□□□□□□□□

```
[postgres@pghost1 ~]$ ll /pgdata/10/data/base/16384/
-rw------- 1 postgres postgres  16384 Nov 28 21:22 3712
...
...
...
-rw------- 1 postgres postgres   8192 Nov 28 21:22 3764_vm
```

□3□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□OID□□□□□□□□1GB□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□OID.<□□□□>□□□□□□□□□□□“OID.<□□□□>”□□□□□□□□□□□□□pg_class□□□relfilenode□□□□□□□□□□□□pg_class□□□□□□□□□□□□□□OID□□relfilenode□□□□□□□VACUUM□TRUNCATE□□□□□relfilenode□□□□□□□

□□□□□

```
mydb=# SELECT oid,relfilenode FROM pg_class WHERE relname =
'tbl';
      oid | relfilenode
----------+-------------
    16387 |       16387
(1 row)
mydb=# \! ls -l /pgdata/10/data/base/16384/16387*
-rw------- 1 postgres postgres 8192 Mar 26 22:22
/pgdata/10/data/base/16384/16387
```

□□□□□□□tbl□□OID□16387□relfilenode□□16387□□□□□□□□“/pgdata/10/data/base/16384/16387”□□□TRUNCATE□□tbl□□□□□□□□□□□□

```
mydb=# TRUNCATE tbl;
TRUNCATE TABLE
mydb=# CHECKPOINT;
CHECKPOINT
mydb=# \! ls -l /pgdata/10/data/base/16384/16387*
ls: cannot access /pgdata/10/data/base/16384/16387*: No such
file or directory
```

□□□□□□□□□tbl□□□□□□□□“/pgdata/10/data/base/16384/16387”□□□□□□□□□□tbl□□□□□□□□□□□

```
postgres@160.40:1922/mydb=# select oid,relfilenode from
pg_class where relname = 'tbl';
      oid | relfilenode
```

```
----------+-------------
    16387 |       24591
(1 row)
postgres@160.40:1922/mydb=# \! ls -l
/pgdata/10/data/base/16384/24591*
-rw------- 1 postgres postgres 0 Apr  2 21:24
/pgdata/10/data/base/16384/24591
```

□□□□□□□□□pg_class□□□tbl□□□□□□□□□“/pgdata/10/data/base/16384/24591”□□□□□□□□<relfilenode>.<□□□>□

□tbl□□□□□□□□□□□□□□□□□□

```
mydb=# insert into tbl (ival,description,created_time)
select (random()*(2*10^9)):: integer as
ival,substr('abcdefghijklmnopqrstuvwxyz',1,
(random()*26)::integer) as
description,date(generate_series(now(), now() + '1 week', '1
day')) as created_time from generate_series(1,2000000);
INSERT 0 16000000
```

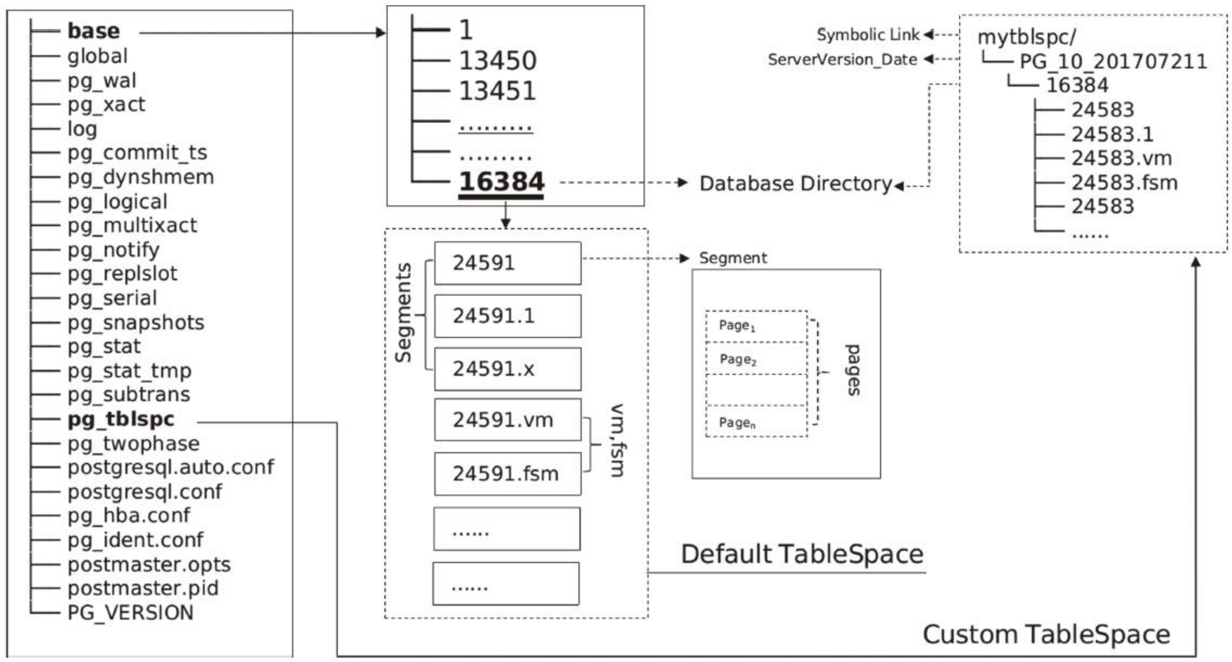
□□□□□□□□□□□□

```
mydb=# SELECT
pg_size_pretty(pg_relation_size('tbl'::regclass));
    pg_size_pretty
----------------
    1068 MB
(1 row)
```

□□□□□□□tbl□□□□□□1068MB□□□□□UPDATE□□□□□□□□□□□□□□□

```
/mydb=# \! ls -lh /pgdata/10/data/base/16384/24591*
-rw------- 1 postgres postgres 1.0G Apr  7 08:44
/pgdata/10/data/base/16384/24591
-rw------- 1 postgres postgres 383M Apr  7 08:44
/pgdata/10/data/base/16384/24591.1
-rw------- 1 postgres postgres 376K Apr  7 08:44
/pgdata/10/data/base/16384/24591_fsm
-rw------- 1 postgres postgres 8.0K Apr  7 08:44
/pgdata/10/data/base/16384/24591_vm
```

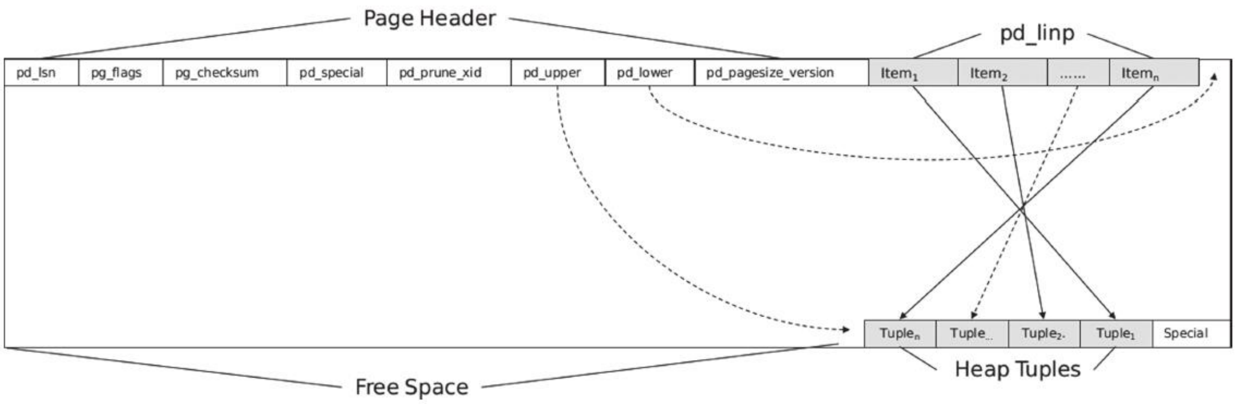
□□□□□□□□□□□□□□<relfilenode>.<□□□>□tbl□□□□□1GB□tbl□□relfilenode□24591□□□1GB□□□□□□□□□□GB□□□□□□□□□□□□□□□24591.1□□□□□□□□□□□□□□□□_fsm□_vm□□□5-2□□□PostgreSQL□□□□□□□□□□□□□□□□



□5-2 PostgreSQL□□□□□□□

□4□□□□□□

□PostgreSQL□□□□□□□□□□Page□□□□□□□□Buffer□□□□□Relation□□□Tuple□□5-3□□□□□□□□Page□□□□□□Page□□□□8kB□□□PostgreSQL□□□BLCKSZ□□□Page□□□□□□□□□BLCKSZ□□□□Page□□□□Page□□□Tuple□□I/O□□□□□□□□□□□□□□□□□□BLCKSZ□□□□□□□□□□□



[illegible]5-3 Page[illegible]

PageHeader[illegible]PageHeader[illegible]

·pd_lsn[illegible]ARIES Recovery Algorithm[illegible]lsn[illegible]PageLSN[illegible]xlog[illegible]LSN[illegible]WAL[illegible]pd_lsn[illegible]xlogid[illegible]64[illegible]LSN[illegible]32[illegible]

·pg_flags[illegible]

·pd_special[illegible]

·pd_lower[illegible]

·pd_upper[illegible]

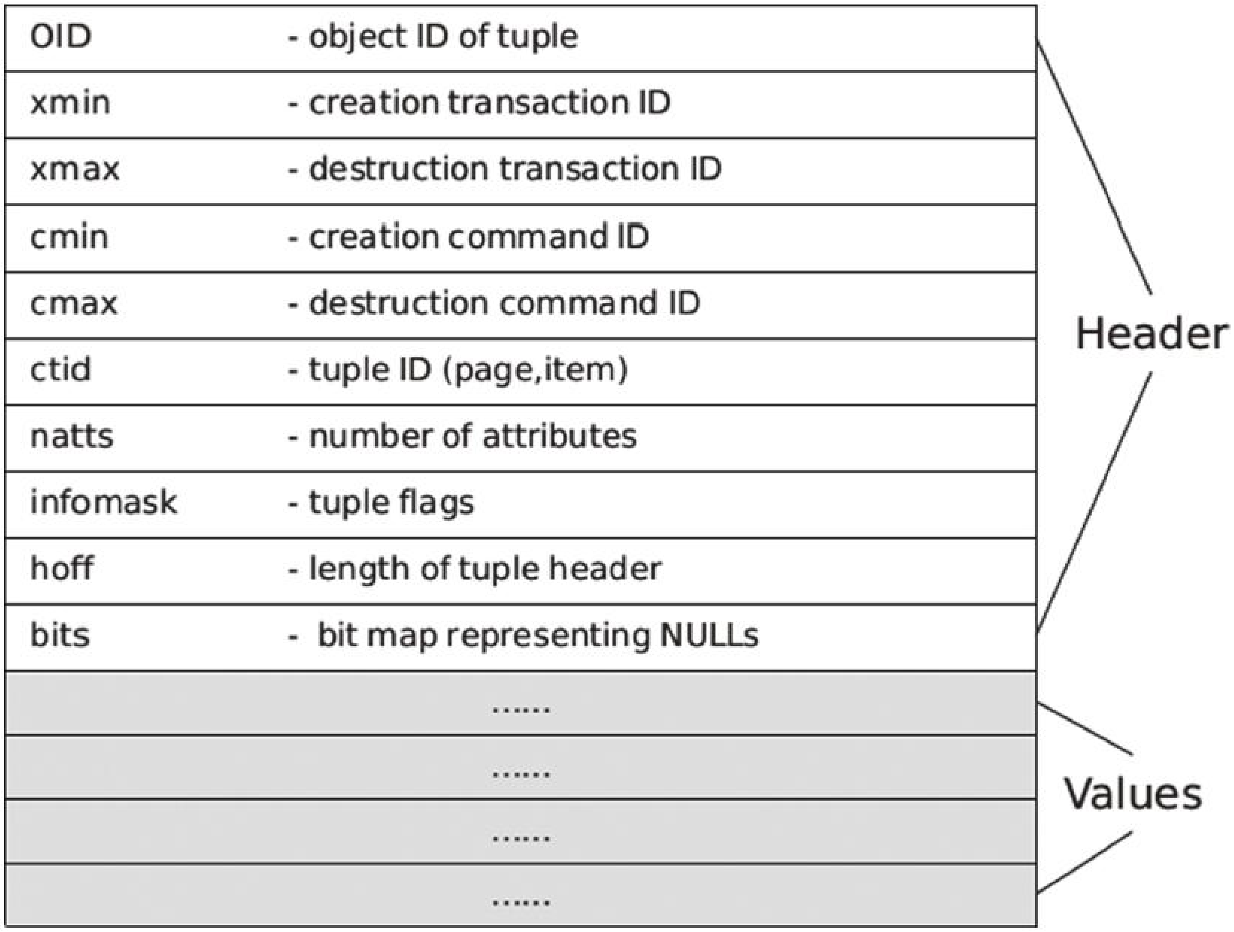
·pd_pagesize_version□□□□PostgreSQL□□□□□□□□□□□□□

·pd_linp[1]□□□□□□□□□5-3□□Item1□Item2□...□Itemn□□□□□□□Tuple□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□pd_lower□□□□□□□□□□□□□□pd_upper□□pd_lower□□□□□□□□□□□pd_upper□□□□□□□□□□□□□□□□□□□□□□□□pd_lsn□pg_checksum□pg_flag□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□B□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□B□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□TID□□□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□TID□□□□□□□□□□

□□Tuple□□□□□□□□□□□□□HeapTupleHeader□□□□□Tuple□□□□□□□5-4□□□□□□Tuple□OID□xmin□cmin□□□□□□□HeapTuple□□□□□Tuple□□□□


OID - object ID of tuple
xmin - creation transaction ID
xmax - destruction transaction ID
cmin - creation command ID
cmax - destruction command ID
ctid - tuple ID (page,item)
natts - number of attributes
infomask - tuple flags
hoff - length of tuple header
bits - bit map representing NULLs
......
......
......
......
Header
Values

□5-4 Tupe□□□□

□5-5□□□□□□□□□□□□□□

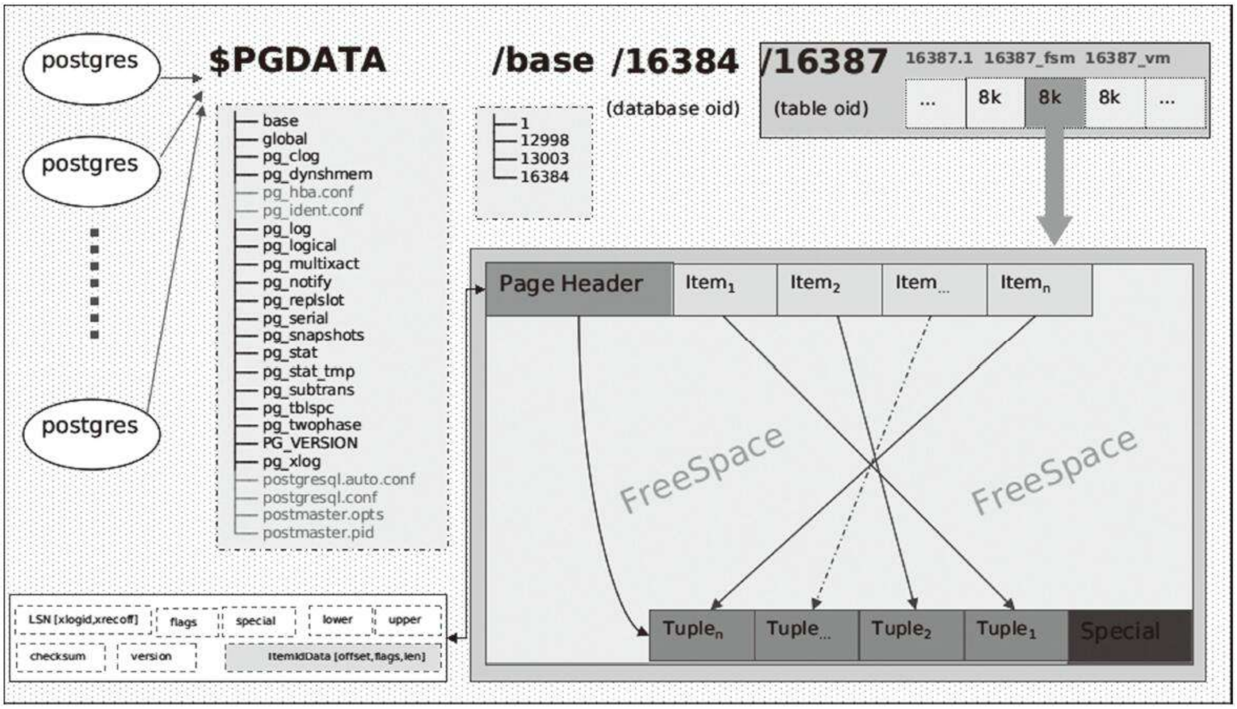

postgres
postgres
postgres
$PGDATA
/base
/16384
(database oid)
/16387
(table oid)
16387.1
16387_fsm
16387_vm
...
8k
8k
8k
...
base
global
pg_clog
pg_dynshmem
pg_hba.conf
pg_ident.conf
pg_log
pg_logical
pg_multixact
pg_notify
pg_replslot
pg_serial
pg_snapshots
pg_stat
pg_stat_tmp
pg_subtrans
pg_tblspc
pg_twophase
PG_VERSION
pg_xlog
postgresql.auto.conf
postgresql.conf
postmaster.opts
postmaster.pid
1
12998
13003
16384
Page Header
Item1
Item2
Item...
Itemn
FreeSpace
FreeSpace
Tuplen
Tuple...
Tuple2
Tuple1
Special
LSN [xlogid,xrecoff]
flags
special
lower
upper
checksum
version
ItemIdData [offset,flags,len]

图5-5 [illegible]

5.2 进程结构

PostgreSQL数据库是由一系列后台/前台进程组成的，这些进程拥有各自的功能，这些进程主要有postmaster主进程、多个postgres进程以及后台的syslogger、checkpointer、bgwriter、walwriter等辅助进程。

5.2.1 主进程和前台进程

主进程postmaster是整个数据库实例postmaster的总控进程，其作用如下：

·启动和停止数据库

·监听客户端连接请求

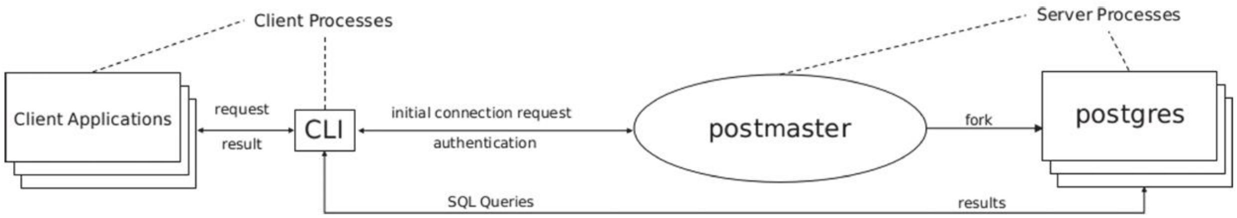
·为每个客户端连接fork单独的postgres服务进程

·服务进程出错时进行修复工作

·管理数据文件

·管理与数据库运行相关的辅助进程

□□□□□□□□□□□□□□□□□□□□postmaster□fork□□□□□□postgres□□□□□□□□□□□postgres□□□□□□□□□□□□□□□□□□postmaster□□□□□□□□postgres□□□□□□□□□□□□5-6□□□



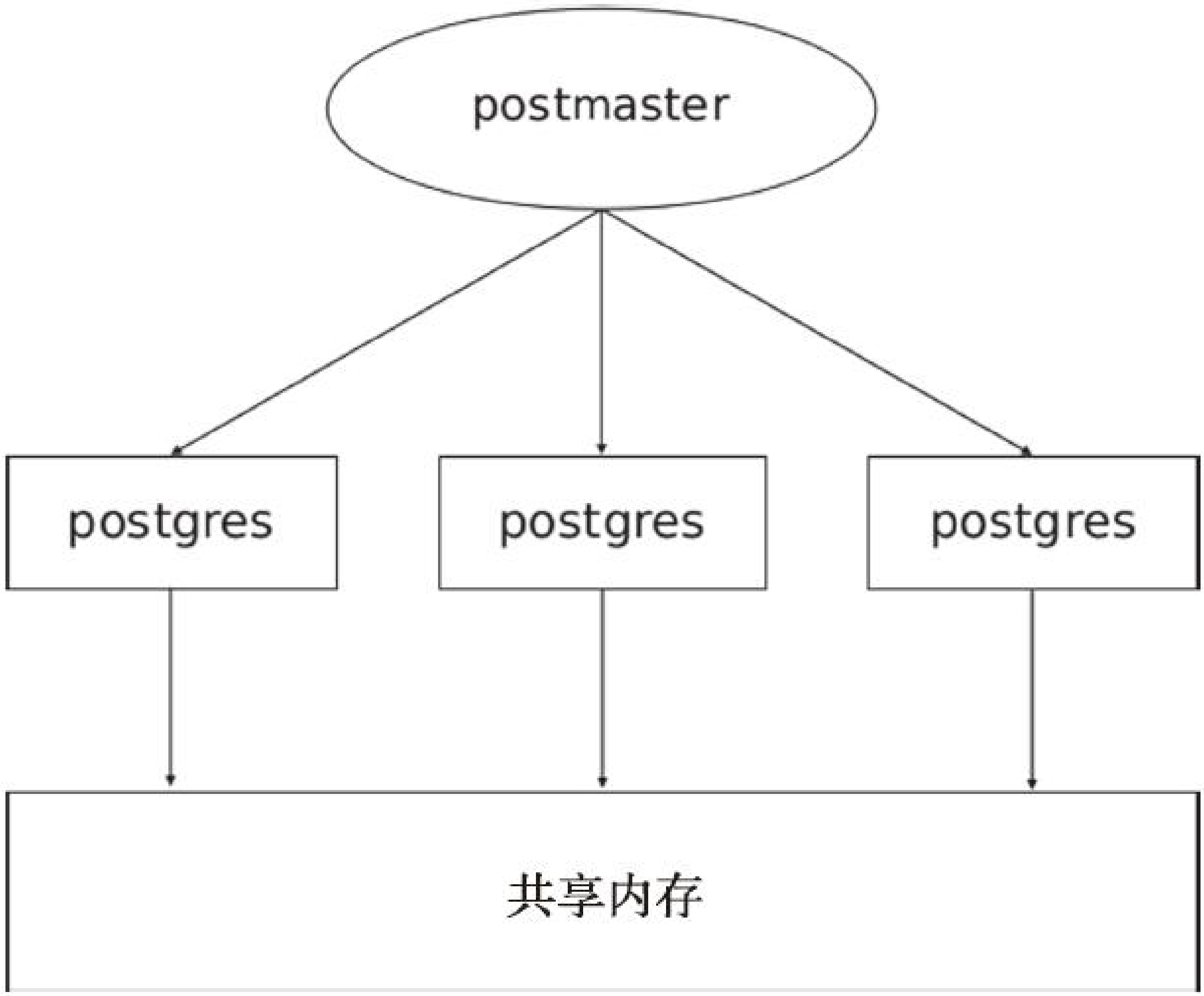
□5-6　□□□□□□□□

PostgreSQL□□□TCP/IP□Unix□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□5-7□□□

5.2.2 □□□

□□□□□postmaster□□□□postgres□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□

·background writer□□□□□bgwriter□□□bgwriter□□□



[illegible]5-7 [illegible]

·autovacuum launcher[illegible]

·WAL writer[illegible]WAL[illegible]WAL[illegible]

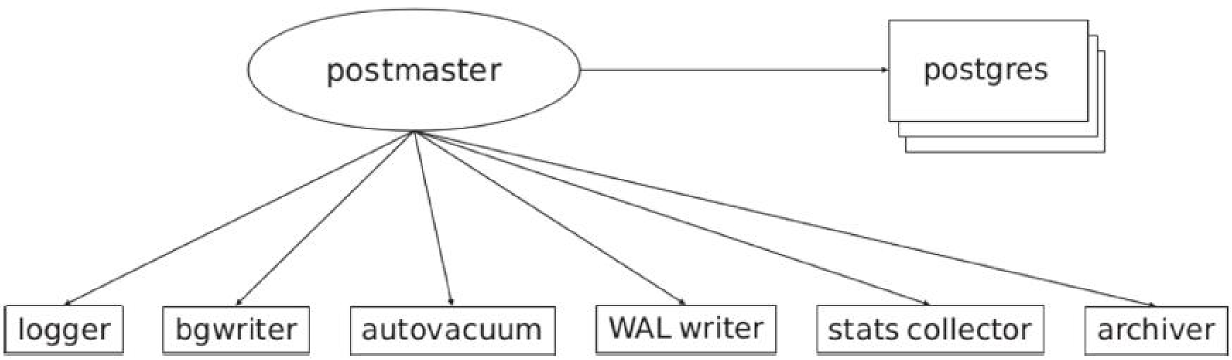
·statistics collector[illegible]

·logging collector[illegible]

·archiver□WAL□□□□□

·checkpointer□□□□□□□

□5-8□□□□□□□□□□□□□□□postmaster□□□□
□□□□



□5-8　□□□□□□□□□□

5.3 [illegible]

PostgreSQL[illegible]

5.3.1 [illegible]

[illegible]
fork[illegible]
work_mem[illegible]maintenance_work_mem[illegible]
temp_buffers[illegible]

·work_mem[illegible]ORDER BY[illegible]DISTINCT[illegible]

·maintenance_work_mem[illegible]
VACUUM[illegible]REINDEX[illegible]CREATE INDEX[illegible]

·temp_buffers[illegible]

5.3.2 [illegible]

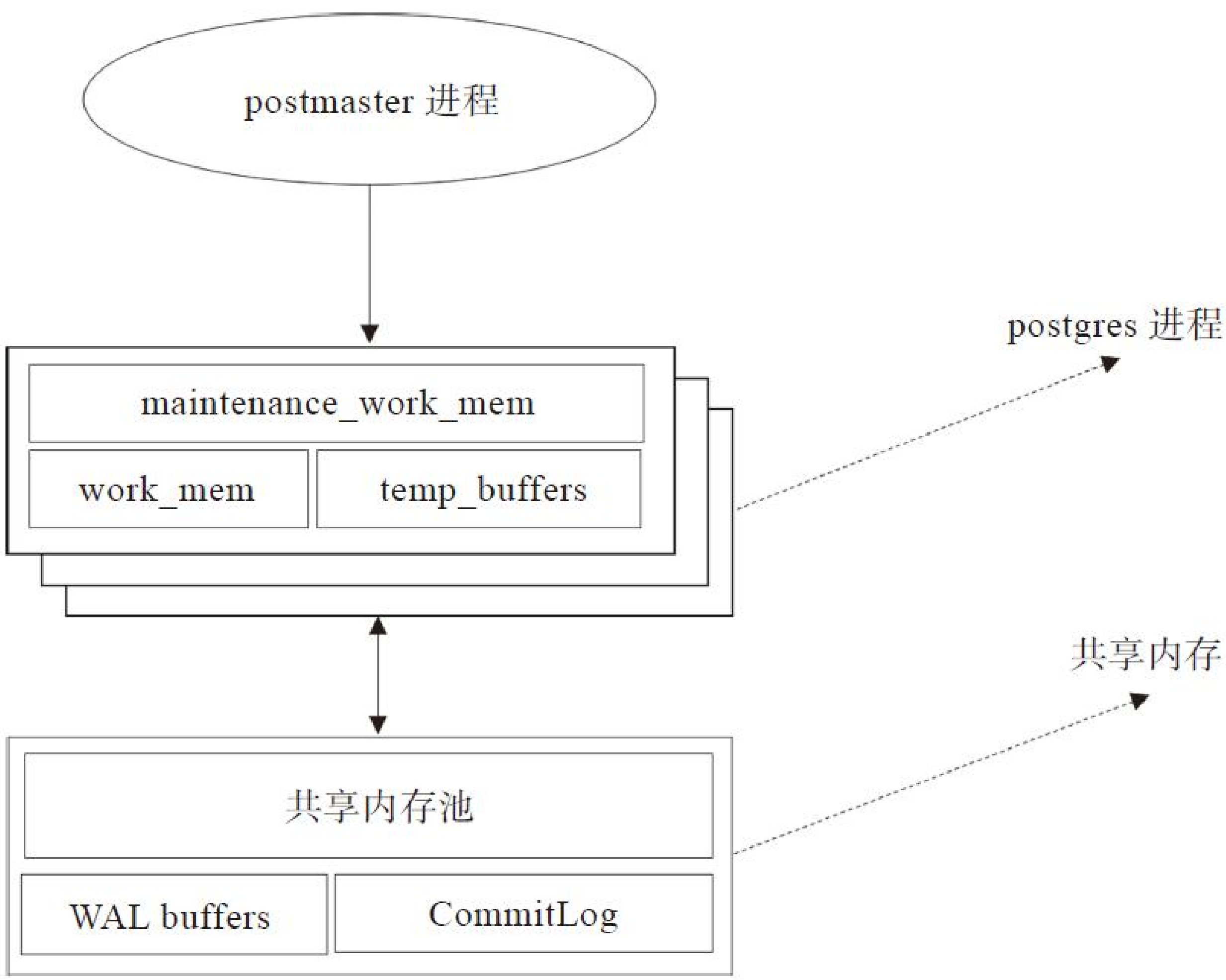
[illegible]PostgreSQL[illegible]

·shared buffer pool□PostgreSQL□□□□□□
□□□□□□□□□□□□□□□□□□□□□□

·WAL buffer□WAL□□□□□□□□□□□□

·CommitLog buffer□PostgreSQL□
Commit Log□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□

□5-9□□□□□□□□□□□



[illegible]5-9 [illegible]

5.4 [illegible]

[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]

□6□　□□□□

□□Oracle□□□□□□□Oracle□□□□□□□□□SELECT□UPDATE□DELETE□□□□□□□□□□□□□□□CPU□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□9.6□□□□□□□□□□□□SQL□□□□□□CPU□□□□□9.6□□□□□□□□□□□□□9.6□□□10□□□□□□□□□□□□□□□□□□□□□□□□□□index-only□□□□□bitmap heap□□□□□□□□□PostgreSQL10□□□□□□□□

6.1 [illegible]

[illegible]PostgreSQL[illegible]

1.max_worker_processes[illegible]integer[illegible]

[illegible]8[illegible]

2.max_parallel_workers[illegible]integer[illegible]

[illegible]8[illegible]max_worker_processes[illegible]max_worker_processes[illegible]

[illegible]max_parallel_workers_per_gather[illegible]

3.max_parallel_workers_per_gather[illegible]integer[illegible]

[illegible]2[illegible]0[illegible]max_worker_processes[illegible]max_parallel_workers[illegible]

□□□□□□□□□□□□□□□□□□□□□□□□□□□CPU□IO□□□□□□□
□□
□□□□□□□□□□□

```
max_worker_processes>max_parallel_workers>max_parallel_worke
rs_per_gather
```

4.parallel_setup_cost□floating point□

□□□□□□□□□□□□□□□□□□□□□1000□

5.parallel_tuple_cost□floating point□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□0.1□

6.min_parallel_table_scan_size□integer□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□8MB□

7.min_parallel_index_scan_size□integer□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□512kb□

8.force_parallel_mode□enum□

□□□□□□□□□□□□□□□□OLTP□□□□□□□□□□□□□□□□□□

□□□□postgresql.conf□□□□□□□□□□□□

```
max_worker_processes = 16
max_parallel_workers_per_gather = 4      # taken from
max_parallel_workers
max_parallel_workers = 8
parallel_tuple_cost = 0.1
parallel_setup_cost = 1000.0
min_parallel_table_scan_size = 8MB
min_parallel_index_scan_size = 512kB
force_parallel_mode = off
```

□□□□□□□□□□□□□□□□□□□□□□4CPU□8GB□□□□□□□

□□ □□□□□□□□□□□□□ max_parallel_workers_per_gather□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□max_parallel_workers□max_worker_processes□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□max_worker_processes□□□□□2□max_parallel_workers_per_gather□□□□□4□□□□□□□□□□□□□□□□□□□□□□□□□□□2□□□□□□□□□□□□□min_parallel_table_scan_size□□□□

□□□□□□□□□□□□□□□□□□□□□□□Workers Planned□□□□□□□□□□□□□Worker Launched□□□□□□□□□□□□□□□

6.2 [illegible]

[illegible]PostgreSQL[illegible]index-only[illegible]bitmap heap[illegible]

6.2.1 [illegible]

[illegible]sequential scan[illegible]CPU[illegible]IO[illegible]OLTP[illegible]

[illegible]5000[illegible]

```
CREATE TABLE test_big1(
id int4,
name character varying(32),
create_time timestamp without time zone default
clock_timestamp());

INSERT INTO test_big1(id,name)
SELECT n, n|| '_test' FROM generate_series(1,50000000) n ;
```

[illegible]

```
mydb=> EXPLAIN SELECT * FROM test_big1 WHERE name='1_test';
            QUERY PLAN
-------------------------------------------------------------
--
    Seq Scan on test_big1  (cost=0.00..991664.00 rows=1
width=25)
        Filter: ((name)::text = '1_test'::text)
(2 rows)
```

[illegible]Seq Scan on test_big1[illegible]test_big1[illegible]PostgreSQL[illegible]9.6[illegible]CPU[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT * FROM test_big1 WHERE
name='1_test';
                                   QUERY PLAN
-------------------------------------------------------------
-------------------
 Gather  (cost=1000.00..523914.10 rows=1 width=25) (actual
time=0.440..1362.675 rows=1 loops=1)
   Workers Planned: 4
   Workers Launched: 4
   ->  Parallel Seq Scan on test_big1  (cost=0.00..522914.00
rows=1 width=25) (actual
       time=1083.280..1355.685 rows=0 loops=5)
         Filter: ((name)::text = '1_test'::text)
         Rows Removed by Filter: 10000000
 Planning time: 0.085 ms
 Execution time: 1367.248 ms
(8 rows)
```

[illegible]Workers Planned[illegible]Worker Launched[illegible]

□□□□□Workers Planned□Worker Launched□□□4□Parallel Seq Scan on test_big1□□□□□□□□□□□□Planning time□□□□□□□□□□□□Execution time□□SQL□□□□□□□□□□□□□□□□□4□□□□SQL□□□□□□□1367□□□□□□□□□□□□□□SQL□□□□□max_parallel_workers_per_gather□□□□□□4□□□□0□□□□□□□□□□□□□□□□□□□0□□□□□□

```
mydb=> SET max_parallel_workers_per_gather =0;
SET
```

□□□□□□□□□□□□□□□

```
mydb=> EXPLAIN ANALYZE SELECT * FROM test_big1 WHERE
name='1_test';
QUERY PLAN
-------------------------------------------------------------
-------------------
 Seq Scan on test_big1  (cost=0.00..991664.00 rows=1
width=25)
      (actual time=0.022. .5329.100 rows=1 loops=1)
    Filter: ((name)::text = '1_test'::text)
    Rows Removed by Filter: 49999999
 Planning time: 0.163 ms
 Execution time: 5329.136 ms
(5 rows)
```

□□□□□□SQL□□□□□5329□□□□□□□□□□□□□□3□□□□

6.2.2 □□□□□□

[illegible]index scan[illegible]

```
mydb=> EXPLAIN SELECT * FROM test_1 WHERE id=1;
                                QUERY PLAN
-------------------------------------------------------------
-------------------
 Index Scan using test_1_pkey on test_1  (cost=0.43..4.45
rows=1 width=26)
    Index Cond: (id = 1)
(2 rows)
```

Index Scan using[illegible]Parallel index scan[illegible]test_big1[illegible]

```
mydb=> CREATE INDEX idx_test_big1_id ON test_big1 USING btree
(id);
CREATE INDEX
```

[illegible]SQL[illegible]ID[illegible]1[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT count(name) FROM test_big1
WHERE id<10000000;
                                    QUERY PLAN
-------------------------------------------------------------
-------------------
Finalize Aggregate  (cost=236183.98..236183.99 rows=1
width=8)
    (actual time=753.392. .753.392 rows=1 loops=1)
    ->  Gather  (cost=236183.96..236183.97 rows=4 width=8)
          (actual time=750. 133..753.384 rows=5 loops=1)
        Workers Planned: 4
        Workers Launched: 4
```

```
       ->  Partial Aggregate  (cost=235183.96..235183.97
rows=1 width=8) (actual time=746.344..746.344 rows=1 loops=5)
              ->  Parallel Index Scan using idx_test_big1_id
on test_big1
                     (cost=0.56. .228921.05 rows=2505162
width=13) (actual tim
                    e=0.029..566.830 rows=2000000 loops=5)
                    Index Cond: (id < 10000000)
 Planning time: 0.116 ms
 Execution time: 762.351 ms
(9 rows)
```

[illegible]4[illegible]762[illegible]

```
mydb=> SET max_parallel_workers_per_gather =0;
SET
```

[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT count(name) FROM test_big1
WHERE id<10000000;
                                 QUERY PLAN
-------------------------------------------------------------
-------------------
Aggregate  (cost=329127.54..329127.55 rows=1 width=8) (actual
time=2636.859..2636.859 rows=1 loops=1)
    ->  Index Scan using idx_test_big1_id on test_big1
(cost=0.56..304075.92 rows= 10020649 width=13) (actual
time=0.031..1654.500 ro
ws=9999999 loops=1)
         Index Cond: (id < 10000000)
 Planning time: 0.132 ms
 Execution time: 2636.920 ms
(5 rows)
```

[illegible]2636[illegible]

[illegible] PostgreSQL10[illegible]SQL[illegible]CPU[illegible]IO[illegible]PostgreSQL 10[illegible]btree[illegible]

6.2.3 [illegible]index-only[illegible]

[illegible]index-only[illegible]index-only[illegible]index-only[illegible]SQL[illegible]SQL[illegible]ID[illegible]100[illegible]

```
mydb=> SET max_parallel_workers_per_gather =0;
SET
```

[illegible]SQL[illegible]

```
mydb=> EXPLAIN SELECT count(*) FROM test_big1 WHERE
id<1000000;
QUERY PLAN
-------------------------------------------------------------
-------------------
    Aggregate  (cost=36060.91..36060.92 rows=1 width=8)
        ->  Index Only Scan using idx_test_big1_id on
```

```
test_big1  (cost=0.56..33313.99 rows=1098767 width=0)
            Index Cond: (id < 1000000)
(3 rows)
```

□□□□□□□□□Index Only Scan□□□□□□ID□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□index-only□□□□□□□□EXPLAIN ANALYZE□□□SQL□□□□□□

```
mydb=> EXPLAIN ANALYZE SELECT count(*) FROM test_big1 WHERE
id<1000000;
                                QUERY PLAN
-------------------------------------------------------------
-------------------
    Aggregate  (cost=35969.89..35969.90 rows=1 width=8)
        (actual time=253.571..253. 571 rows=1 loops=1)
        ->  Index Only Scan using idx_test_big1_id on
test_big1  (cost=0.56..33232.22 rows=1095066 width=0) (actual
time=0.038..179.005 rows=999999 loops=1)
            Index Cond: (id < 1000000)
            Heap Fetches: 999999
    Planning time: 0.103 ms
    Execution time: 253.617 ms
(6 rows)
```

□□□□□253□□□index-only□□□□□□□□□□□index-only□□□□□□□□□index-only□□□□□□□□□□□□□□□□□□□□

```
mydb=> SET max_parallel_workers_per_gather TO default;
SET
```

□□□□□□□□□□□□□□□

```
mydb=> EXPLAIN ANALYZE SELECT count(*) FROM test_big1 WHERE
id<1000000;
                            QUERY PLAN
-------------------------------------------------------------
-------------------
    Finalize Aggregate  (cost=26703.66..26703.67 rows=1
width=8)
        (actual time=81. 285..81.285 rows=1 loops=1)
        ->  Gather  (cost=26703.64..26703.65 rows=4 width=8)
              (actual time=81. 121..81.277 rows=5 loops=1)
            Workers Planned: 4
            Workers Launched: 4
            ->  Partial Aggregate  (cost=25703.64..25703.65
rows=1 width=8) (actual time=75.778..75.778 rows=1 loops=5)
                ->  Parallel Index Only Scan using
idx_test_big1_id on test_big1  (cost=0.56..25019.22
rows=273766 width=0)
                      (actual time=0.045..59.398 rows=200000
loops=5)
                    Index Cond: (id < 1000000)
                    Heap Fetches: 183366
    Planning time: 0.113 ms
    Execution time: 83.364 ms
(10 rows)
```

□□□□□□□□□Parallel Index Only Scan□□□□□□□□index-only□□□□□□□□□83□□□□□□□□□□□□□□□□

6.2.4 □□bitmap heap□□

□□□□bitmap heap□□□□□□□□Bitmap Index□□□Bitmap Heap□□□□SQL□where□□□□

□or□□□□□□□Bitmap Index□□□□□□□□

```
mydb=> EXPLAIN SELECT * FROM test_big1 WHERE id=1 OR id=2;
                                     QUERY PLAN
-------------------------------------------------------------
-------------------
Bitmap Heap Scan on test_big1  (cost=5.15..9.17 rows=2
width=25)
    Recheck Cond: ((id = 1) OR (id = 2))
    ->  BitmapOr  (cost=5.15..5.15 rows=2 width=0)
        ->  Bitmap Index Scan on idx_test_big1_id
(cost=0.00..2.57 rows=1 width=0)
            Index Cond: (id = 1)
        ->  Bitmap Index Scan on idx_test_big1_id
(cost=0.00..2.57 rows=1 width=0)
            Index Cond: (id = 2)
(7 rows)
```

□□□□□□□□□□□□□□□□Bitmap Index□□□□□□□□□□□Bitmap Index□□□□□□□□□□□□□□□□□□□test_big1□□□□Bitmap Heap□□□Bitmap Heap□□□□□□□□□□□□SQL□□□□□□□□ID□□□□□□□□

```
EXPLAIN ANALYZE SELECT count(*) FROM test_big1 WHERE id
<1000000 OR id > 49000000;
```

□□□□□□□□□

```
mydb=> EXPLAIN ANALYZE SELECT count(*) FROM test_big1 WHERE
id <1000000 OR id > 49000000;
                                QUERY PLAN
-------------------------------------------------------------
-------------------
Finalize Aggregate  (cost=406220.88..406220.89 rows=1
```

```
width=8)
    (actual time=241. 186..241.186 rows=1 loops=1)
    ->  Gather  (cost=406220.46..406220.87 rows=4 width=8)
          (actual time=241. 033..241.174 rows=5 loops=1)
        Workers Planned: 4
        Workers Launched: 4
        ->  Partial Aggregate  (cost=405220.46..405220.47
rows=1 width=8) (actual time=237.266..237.266 rows=1 loops=5)
            ->  Parallel Bitmap Heap Scan on test_big1
(cost=28053.14..403933.27 rows=514876 width=0) (actual
time=83.059..189.073 rows=400000 loops=5)
                Recheck Cond: ((id < 1000000) OR (id >
49000000))
                Heap Blocks: exact=2982
                ->  BitmapOr  (cost=28053.14..28053.14
rows=2081101 width=0) (actual time=83.657..83.657 rows=0
loops=1)
                    ->  Bitmap Index Scan on idx_test_big1_id
(cost=0.00.. 14219.03 rows=1094996 width=0)
                          (actual time=42.187..42.187
rows=999999 loops=1)
                        Index Cond: (id < 1000000)
                    ->  Bitmap Index Scan on idx_test_big1_id
                          (cost=0.00.. 12804.35 rows=986105
width=0)
                           (actual time=41.467..41.467
rows=1000000 loops=1)
                        Index Cond: (id > 49000000)
 Planning time: 0.157 ms
 Execution time: 242.824 ms
(15 rows)
```

[illegible]Bitmap Heap[illegible]4[illegible]242ms[illegible]

```
mydb=> SET max_parallel_workers_per_gather =0;
SET
```

[illegible]SQL[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT count(*) FROM test_big1 WHERE
id <1000000 OR id > 49000000;
                                    QUERY PLAN
-------------------------------------------------------------
-------------------
Aggregate  (cost=432494.42..432494.43 rows=1 width=8) (actual
time=466.273.. 466.273 rows=1 loops=1)
    ->  Bitmap Heap Scan on test_big1
(cost=28053.14..427345.66 rows=2059505 width=0) (actual
time=86.859..299.822 rows=1999999 loops=1)
        Recheck Cond: ((id < 1000000) OR (id > 49000000))
        Heap Blocks: exact=13724
        ->  BitmapOr  (cost=28053.14..28053.14 rows=2081101
width=0)
              (actual time=84.782..84.782 rows=0 loops=1)
            ->  Bitmap Index Scan on idx_test_big1_id
(cost=0.00..14219.03 rows= 1094996 width=0) (actual
time=42.704..42.704 rows=999999 loops=1)
                Index Cond: (id < 1000000)
            ->  Bitmap Index Scan on idx_test_big1_id
(cost=0.00..12804.35 rows= 986105 width=0) (actual
time=42.076..42.076 rows=1000000 loops=1)
                Index Cond: (id > 49000000)
 Planning time: 0.152 ms
 Execution time: 466.323 ms
(11 rows)
```

[illegible]Bitmap Heap[illegible]466[illegible]

6.3 □□□□

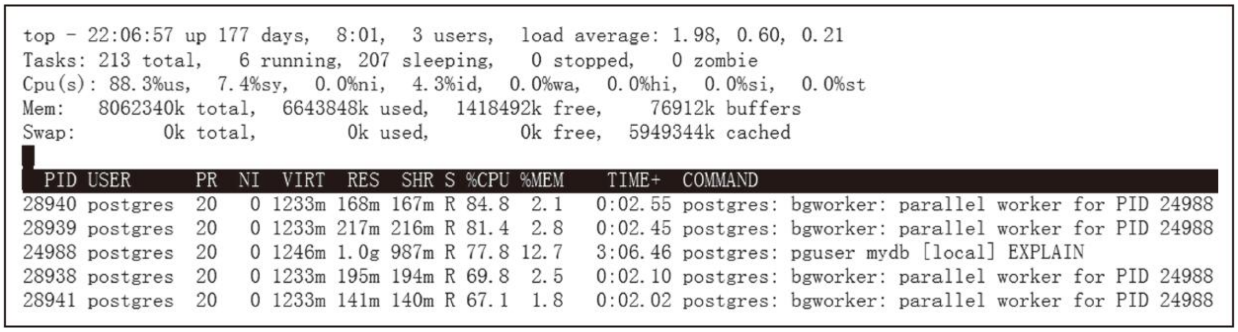
□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□count□□□sum□□□□□□□□SQL□□□□□count□□□□□□□□□□□□□□□□□□□□□

```
mydb=> EXPLAIN ANALYZE SELECT count(*) FROM test_big1;
QUERY PLAN
------------------------------------------------------------
--------------------
Finalize Aggregate  (cost=525335.91..525335.92 rows=1
width=8)
    (actual time=2468.593.. 2468.594 rows=1 loops=1)
    ->  Gather  (cost=525335.89..525335.90 rows=4 width=8)
          (actual time=2468. 386..2468.585 rows=5 loops=1)
        Workers Planned: 4
        Workers Launched: 4
        ->  Partial Aggregate  (cost=524335.89..524335.90
rows=1 width=8)
              (actual time=2463.532..2463.532 rows=1
loops=5)
            ->  Parallel Seq Scan on test_big1
(cost=0.00..493083.91
                  rows= 12500791 width=0) (actual
time=0.019..1456.506 rows=10000000
                  loops=5)
 Planning time: 0.089 ms
 Execution time: 2474.970 ms
(8 rows)
```

□□□□□□□□□□□□□□Partial Aggregate□□□□□□□□□□□□□□□Finalize Aggregate□□SQL□□□□□2474□□□□□□□□□□□□top□□□□□EXPLAIN ANALYZE□□□□□□□□□□□6-1□

[illegible]count[illegible]

```
mydb=> SET max_parallel_workers_per_gather = 0 ;
SET
```



[illegible]6-1 top[illegible]

[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT count(*) FROM test_big1;
QUERY PLAN
------------------------------------------------------------
--------------------
Aggregate  (cost=993115.55..993115.56 rows=1 width=8)
    (actual time=8655.614.. 8655.615 rows=1 loops=1)
    ->  Seq Scan on test_big1  (cost=0.00..868107.64
rows=50003164 width=0) (actual time=0.019..4898.227
rows=50000000 loops=1)
 Planning time: 0.106 ms
 Execution time: 8655.666 ms
(4 rows)
```

[illegible]2474[illegible]8865[illegible]3[illegible]

2[illegible]

```
mydb=> SET max_parallel_workers_per_gather =2;
SET
```

[illegible]SQL[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT count(*) FROM test_big1;
                                QUERY PLAN
------------------------------------------------------------
--------------------
Finalize Aggregate  (cost=629509.16..629509.17 rows=1
width=8)
    (actual time=3305. 585..3305.585 rows=1 loops=1)
    ->  Gather  (cost=629509.15..629509.16 rows=2 width=8)
          (actual time=3305. 512..3305.578 rows=3 loops=1)
        Workers Planned: 2
        Workers Launched: 2
        ->  Partial Aggregate  (cost=628509.15..628509.16
rows=1 width=8) (actual time=3302.362..3302.362 rows=1
loops=3)
            ->  Parallel Seq Scan on test_big1
(cost=0.00..576422.52 rows= 20834652 width=0) (actual
time=0.021..2000.427 rows=16666667 loops=3)
 Planning time: 0.112 ms
 Execution time: 3314.371 ms
(8 rows)
```

[illegible]3314[illegible]6[illegible]8[illegible]SQL[illegible]6-1[illegible]

[illegible]6-1 [illegible]

并行进程数	Count() 执行时间
0	8865 毫秒
2	3314 毫秒
4	2474 毫秒
6	2479 毫秒
8	2446 毫秒

[illegible]6-1[illegible]8[illegible]4[illegible]6[illegible]4[illegible]CPU[illegible]4[illegible]CPU[illegible]

sum[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT sum(hashtext(name)) FROM
test_big1;
                                     QUERY PLAN
------------------------------------------------------------
--------------------
Finalize Aggregate  (cost=555163.25..555163.26 rows=1
width=8)
    (actual time=3307. 939..3307.939 rows=1 loops=1)
    ->  Gather  (cost=555162.83..555163.24 rows=4 width=8)
          (actual time=3307. 694..3307.934 rows=5 loops=1)
        Workers Planned: 4
        Workers Launched: 4
        ->  Partial Aggregate  (cost=554162.83..554162.84
rows=1 width=8) (actual time=3303.628..3303.628 rows=1
loops=5)
            >  Parallel Seq Scan on test_big1
(cost=0.00..491663.22 rows=12499922
               width=13) (actual time=0.045..1837.554
rows=10000000 loops=5)
 Planning time: 0.078 ms
 Execution time: 3308.999 ms
(8 rows)
```

min[illegible]max[illegible]

6.4 [illegible]

[illegible]nested loop[illegible]merge join[illegible]hash join[illegible]

6.4.1 Nested loop[illegible]

[illegible]Nested loop[illegible]

```
for (i = 0; i < length(outer); i++)
    for (j = 0; j < length(inner); j++)
        if (outer[i] == inner[j])
            output(outer[i], inner[j]);
```

[illegible]Nested loop[illegible]test_small[illegible]

```
CREATE TABLE test_small(id int4, name character
varying(32));

INSERT INTO test_small(id,name)
SELECT n, n|| '_small' FROM generate_series(1,8000000) n ;
```

[illegible]

```
mydb=> CREATE INDEX idx_test_small_id ON test_small USING
btree (id);
CREATE INDEX

mydb=> ANALYZE test_small;
ANALYZE
```

ANALYZE[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT test_small.name
        FROM test_big1, test_small
        WHERE test_big1.id = test_small.id
        AND test_small.id < 10000;
QUERY PLAN
------------------------------------------------------------
--------------------
Gather  (cost=1138.18..31628.76 rows=10217 width=13) (actual
time=1.036..16.207 rows=9999 loops=1)
    Workers Planned: 3
    Workers Launched: 3
    ->  Nested Loop  (cost=138.18..29607.06 rows=3296
width=13)
          (actual time= 0.244..10.895 rows=2500 loops=4)
        ->  Parallel Bitmap Heap Scan on test_small
(cost=137.61..14796.65
              rows= 3296 width=17) (actual time=0.203..0.653
rows=2500 loops=4)
                Recheck Cond: (id < 10000)
                Heap Blocks: exact=10
                ->  Bitmap Index Scan on idx_test_small_id
(cost=0.00..135.06 rows=10217 width=0) (actual
time=0.652..0.652 rows=9999 loops=1)
                    Index Cond: (id < 10000)
        ->  Index Only Scan using idx_test_big1_id on
test_big1  (cost=0.56..4.48 rows=1 width=4) (actual
```

```
time=0.003..0.003 rows=1 loops=9999)
                    Index Cond: (id = test_small.id)
                    Heap Fetches: 1674
 Planning time: 0.427 ms
 Execution time: 17.548 ms
(14 rows)
```

[illegible]test_big1[illegible]Index Only[illegible]id[illegible]10000[illegible]Nested loop[illegible]test_small1[illegible]Bitmap Heap[illegible]id[illegible]10000[illegible]SQL[illegible]17.5[illegible]

```
mydb=> SET max_parallel_workers_per_gather =0;
SET
```

[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT test_small.name
        FROM test_big1, test_small
        WHERE test_big1.id = test_small.id
        AND test_small.id < 10000;
                                QUERY PLAN
-----------------------------------------------------------------
-------------
Nested Loop  (cost=1.00..46213.80 rows=10217 width=13)
(actual time=0.025..29.054 rows=9999 loops=1)
    ->  Index Scan using idx_test_small_id on test_small
(cost=0.43..304.23 rows=10217 width=17) (actual
time=0.014..2.284 rows=9999 loops=1)
            Index Cond: (id < 10000)
    ->  Index Only Scan using idx_test_big1_id on test_big1
(cost=0.56..4.48 rows=1 width=4) (actual time=0.002..0.002
rows=1 loops=9999)
```

```
            Index Cond: (id = test_small.id)
            Heap Fetches: 9999
 Planning time: 0.262 ms
 Execution time: 29.606 ms
(8 rows)
```

[illegible]SQL[illegible]29.6[illegible]

6.4.2 Merge join[illegible]

Merge join[illegible]Merge join[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT test_small.name
        FROM test_big1, test_small
        WHERE test_big1.id = test_small.id
        AND test_small.id < 200000;
                                         QUERY PLAN
------------------------------------------------------------
---------------------
 Gather  (cost=1001.79..195146.04 rows=204565 width=13)
     (actual time=0.308.. 160.192 rows=199999 loops=1)
    Workers Planned: 4
    Workers Launched: 4
    ->  Merge Join  (cost=1.79..173689.54 rows=51141
width=13)
          (actual time=4.059.. 127.164 rows=40000 loops=5)
            Merge Cond: (test_big1.id = test_small.id)
            ->  Parallel Index Only Scan using
idx_test_big1_id on test_big1  (cost=0.56..1017275.56
rows=12500000 width=4)
                  (actual time=0.044..16.223 rows=40001
loops=5)
                Heap Fetches: 50875
            ->  Index Scan using idx_test_small_id on
```

```
test_small  (cost=0.43..6010.32
                 rows=204565 width=17) (actual
time=0.034..64.427 rows=199999 loops=5)
                Index Cond: (id < 200000)
 Planning time: 0.255 ms
 Execution time: 171.755 ms
(11 rows)
```

□□□□□□□□□□□□□171□□□□□□□□□□□□□□□□□□□□□

```
mydb=> SET max_parallel_workers_per_gather =0;
SET
```

□□□□□□□□□□□□□□

```
mydb=> EXPLAIN ANALYZE SELECT test_small.name
        FROM test_big1, test_small
        WHERE test_big1.id = test_small.id
        AND test_small.id < 200000;
                                 QUERY PLAN
------------------------------------------------------------
---------------------
 Merge Join  (cost=1.79..249728.34 rows=204565 width=13)
(actual time=0.034.. 198.331 rows=199999 loops=1)
    Merge Cond: (test_big1.id = test_small.id)
    ->  Index Only Scan using idx_test_big1_id on test_big1
(cost=0.56..1392275.56 rows=50000000 width=4) (actual
time=0.018..60.445
          rows=200000 loops=1)
            Heap Fetches: 200000
    ->  Index Scan using idx_test_small_id on test_small
(cost=0.43..6010.32 rows=204565 width=17) (actual
time=0.013..50.531 rows=199999 loops=1)
            Index Cond: (id < 200000)
 Planning time: 0.264 ms
```

```
 Execution time: 209.387 ms
(8 rows)
```

[illegible]SQL[illegible]209[illegible]

6.4.3 Hash join[illegible]

PostgreSQL[illegible]Hash join[illegible]Hash join[illegible]Hash join[illegible]

```
mydb=> DROP INDEX idx_test_big1_id;
DROP INDEX
mydb=> DROP INDEX idx_test_small_id ;
DROP INDEX

mydb=> SET max_parallel_workers_per_gather =0;
SET
```

[illegible]

```
mydb=> EXPLAIN SELECT test_small.name
        FROM test_big1 JOIN test_small ON (test_big1.id =
test_small.id)
        AND test_small.id < 100;
                                   QUERY PLAN
------------------------------------------------------------
---------------------
 Hash Join  (cost=150871.78..1205043.77 rows=800 width=13)
    Hash Cond: (test_big1.id = test_small.id)
    ->  Seq Scan on test_big1  (cost=0.00..866664.00
```

```
rows=50000000 width=4)
    ->  Hash  (cost=150861.78..150861.78 rows=800 width=17)
            ->  Seq Scan on test_small
(cost=0.00..150861.78 rows=800 width=17)
                Filter: (id < 100)
(6 rows)
```

□□□□□□□SQL□□□□□□□□□□□

```
mydb=> EXPLAIN ANALYZE SELECT test_small.name
        FROM test_big1 JOIN test_small ON (test_big1.id =
test_small.id)
        AND test_small.id < 100;
                               QUERY PLAN
------------------------------------------------------------
---------------------
 Hash Join  (cost=151317.06..1206944.59 rows=802 width=13)
    (actual time=735. 778..11259.744 rows=99 loops=1)
    Hash Cond: (test_big1.id = test_small.id)
    ->  Seq Scan on test_big1  (cost=0.00..868107.64
rows=50003164 width=4) (actual time=0.015..5424.515
rows=50000000 loops=1)
    ->  Hash  (cost=151307.04..151307.04 rows=802 width=17)
          (actual time=735. 750..735.750 rows=99 loops=1)
            Buckets: 1024  Batches: 1  Memory Usage: 13kB
            ->  Seq Scan on test_small
(cost=0.00..151307.04 rows=802 width=17)
                (actual time=0.010..735.731 rows=99 loops=1)
                Filter: (id < 100)
                Rows Removed by Filter: 7999901
 Planning time: 0.134 ms
 Execution time: 11259.789 ms
(10 rows)
```

□□□□□□□□□□□11259□□□□□□□4□□□□□□□
□□

```
mydb=> SET max_parallel_workers_per_gather =4;
SET
```

□□□□SQL□□□□□□

```
mydb=> EXPLAIN ANALYZE SELECT test_small.name
        FROM test_big1 JOIN test_small ON (test_big1.id =
test_small.id)
        AND test_small.id < 100;
                                  QUERY PLAN
------------------------------------------------------------
---------------------
 Gather  (cost=152317.06..692280.94 rows=802 width=13)
     (actual time=1152. 263..4304.757 rows=99 loops=1)
    Workers Planned: 4
    Workers Launched: 4
    ->  Hash Join  (cost=151317.06..691280.94 rows=200
width=13) (actual time= 3667.973..4298.423 rows=20 loops=5)
            Hash Cond: (test_big1.id = test_small.id)
            ->  Parallel Seq Scan on test_big1
(cost=0.00..493083.91 rows=12500791 width=4) (actual
time=0.025..1737.558 rows=10000000 loops=5)
            ->  Hash  (cost=151307.04..151307.04 rows=802
width=17)
                  (actual time=1106.665..1106.665 rows=99
loops=5)
                Buckets: 1024  Batches: 1  Memory Usage:
13kB
                ->  Seq Scan on test_small
(cost=0.00..151307.04 rows=802 width=17)
                     (actual time=863.670..1106.624 rows=99
loops=5)
                    Filter: (id < 100)
                    Rows Removed by Filter: 7999901
 Planning time: 0.156 ms
 Execution time: 4315.928 ms
(13 rows)
```

□□□□□□□□□□□□□□4□□□□□□□□□□□□4315□□□

6.5 □□□□

□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□index-only□□□□□bitmap heap□□SQL□□□□□□□□□□□□□□□□□3□4□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

[illegible]7[illegible] [illegible]

[illegible]Lock-Based Concurrency Control[illegible]MVCC[illegible]Multi-Version Concurrency Control[illegible]PostgreSQL[illegible]PostgreSQL MVCC[illegible]

7.1 [illegible]

7.1.1 [illegible]

[illegible]PostgreSQL[illegible]BEGIN...END/COMMIT/ROLLBACK[illegible]BEGIN...END/COMMIT/ROLLBACK[illegible]

·[illegible]Atomicity[illegible]

·[illegible]Consistency[illegible]

·[illegible]Isolation[illegible]

·[illegible]Durability[illegible]

[illegible]ACID[illegible]WAL[illegible]MVCC[illegible]

7.1.2 [illegible]

[illegible]PostgreSQL[illegible]（Dirty read[illegible]Non-repeatable read[illegible]Phantom Read[illegible]（Serialization Anomaly[illegible]

[illegible] [illegible]PostgreSQL[illegible]READ UNCOMMITTED[illegible]READ COMMITTED[illegible]PostgreSQL[illegible]PostgreSQL[illegible]READ COMMITTED[illegible]

□□PostgreSQL□□□□□□□□□READ COMMITTED □□□□□□□□□□□□□□□□□□□□

1.□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□INSERT□UPDATE□DELETE□□□□□□□□□□□□□ROLLBACK□□

□□□□□□□□□□□□□□□□□□□□□□

```
CREATE TABLE tbl_mvcc (
    id INT NOT NULL AUTO_INCREMENT,
    ival INT,
    PRIMARY KEY (id)
);
-- □□□□□□□□
INSERT INTO tbl_mvcc (ival) VALUES (1);
```

□□□□□□□□□□□□□T1□T2□□□□□

T1	T2
set session transaction isolation level read uncommitted;start transaction; select * from tbl_mvcc where id = 1; /* id=1,ival=1 */	

（续）

T1	T2
	start transaction; update tbl_mvcc set ival = 10 where id = 1;
select * from tbl_mvcc where id = 1; /* id=1,ival=10 */	
	rollback;

[illegible]T1[illegible]tbl_mvcc[illegible]id=1[illegible]ival=1[illegible]T2[illegible]id=1[illegible]ival[illegible]10[illegible]T1[illegible]tbl_mvcc[illegible]T2[illegible]ival[illegible]1[illegible]T1[illegible]ival[illegible]10[illegible]T2[illegible]ROLLBACK[illegible]T1[illegible]

2.[illegible]

[illegible]

[illegible]

```
CREATE TABLE tbl_mvcc (
    id SERIAL PRIMARY KEY,
    ival INT
);
-- [illegible]
INSERT INTO tbl_mvcc (ival) VALUES (1);
```

[illegible]T1[illegible]T2[illegible]

T1	T2
BEGIN TRANSACTION ISOLATION LEVEL READ COMMITTED;	
SELECT id,ival FROM tbl_mvcc WHERE id = 1;	
id | ival ---+------ 1 | 1 (1 row)	
	BEGIN; UPDATE tbl_mvcc SET ival = 10 WHERE id = 1; COMMIT;
SELECT id,ival FROM tbl_mvcc WHERE id = 1; id | ival ---+------ 1 | 10 (1 row)	
END;	

[illegible]T1[illegible]tbl_mvcc[illegible]id=1[illegible]ival=1[illegible]T2[illegible]id=1[illegible]ival[illegible]10[illegible]T2[illegible]COMMIT[illegible]T1[illegible]tbl_mvcc[illegible]ival[illegible]10[illegible]SELECT[illegible]T2[illegible]T2[illegible]T1[illegible]

3.[illegible]

[illegible]

□□□□□□□□□□□□□□□□□□□□INSERT□DELETE□□□□□□□□□□□□□□□□

　　□□□□□□□□□□□□□□□□□□□□□□tbl_mvcc□□□□□□□□□□□□□□

T1	T2
BEGIN TRANSACTION ISOLATION LEVEL READ COMMITTED; SELECT id,ival FROM tbl_mvcc WHERE id > 3 AND id < 10; id | ival ---+------ 4 | 4 5 | 5 (2 rows)	

	BEGIN;
	INSERT INTO tbl_mvcc (id , ival) VALUES (6 , 6);
	END;

SELECT id,ival FROM tbl_mvcc WHERE id > 3 AND id < 10; id | ival ---+------ 4 | 4 5 | 5 6 | 6 (3 rows)	
END;	

　　□□□□□□□□□□T1□tbl_mvcc□□□□□□□id□□3□□□10□□□□□□□□□□□□□□□T2□□□□□□□□id□□6□□□□□□□□□□□□□□T1□WHERE□□□□id□□3□□□□10□□□□□□□□T1□□UPDATE□□□□□□□□□□□□□□INSERT□DELETE□□□□□□

7.1.3 ANSI SQL[illegible]

[illegible]ANSI SQL[illegible]

·Read Uncommitted[illegible]

·Read Committed[illegible]
PostgreSQL[illegible]

·Repeatable Read[illegible]

·Serializable[illegible]

[illegible]ANSI SQL[illegible]

隔离级别	脏读	不可重复读	幻读
Read Uncommitted	可能	可能	可能
Read Committed	不可能	可能	可能
Repeatable Read	不可能	不可能	可能
Serializable	不可能	不可能	不可能

[illegible]Read Committed[illegible]

7.2 PostgreSQL[illegible]

[illegible]SQL[illegible]“[illegible]”[illegible]SQL[illegible]Read Uncommitted[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]Read Uncommitted[illegible]Read Committed[illegible]PostgreSQL[illegible]Repeatable Read[illegible]SQL[illegible]SQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]T1[illegible]T2[illegible]

T1	T2

`BEGIN TRANSACTION ISOLATION LEVEL REPEATABLE READ ;`
`SELECT id,ival FROM tbl_mvcc WHERE id = 1;`
`id | ival`
`---+------`
` 1 | 1`
`(1 rows)`	

	`UPDATE tbl_mvcc SET ival = ival * 10 WHERE id = 1;`
`UPDATE tbl_mvcc SET ival = ival + 1 WHERE id = 1;` `ERROR: could not serialize access due to concurrent update`	
`ROLLBACK`	

[illegible]T1[illegible]id=1[illegible]T2[illegible]T1[illegible]T1[illegible]T1[illegible]T2[illegible]T1[illegible]T1[illegible]T2[illegible]ival=10[illegible]T1[illegible]UPDATE[illegible]T1[illegible]ival=1[illegible]Serializable[illegible]Repeatable Read[illegible]PostgreSQL[illegible]Serializable[illegible]

[illegible]PostgreSQL[illegible]

隔离级别	脏读	不可重复读	幻读	序列化异常
Read Uncommitted	不可能	可能	可能	可能
Read Committed	不可能	可能	可能	可能
Repeatable Read	不可能	不可能	不可能	可能
Serializable	不可能	不可能	不可能	不可能

□□□□□□□□□□□□□Repeatable Read□□□□□
□□□□□□□□□□□□□□

T1	T2
BEGIN TRANSACTION ISOLATION LEVEL REPEATABLE READ; SELECT id,ival FROM tbl_mvcc WHERE id > 3 AND id < 10; id | ival ---+------ 4 | 4 5 | 5 (2 rows)	

	BEGIN;
	INSERT INTO tbl_mvcc (id , ival) VALUES (6 , 6);
	END;

SELECT id,ival FROM tbl_mvcc WHERE id > 3 AND id < 10; id | ival ---+------ 4 | 4 5 | 5 (2 rows)	
END;	

□□□□□□□□□□T1□tbl_mvcc□□□□□□□id□□3□□□□10□□□□□□□□□□□□□□□T2□□□□□□□□id□□6□□□□□□□□□□□□□□T1□WHERE□□□□id□□3□□□□10□□□□□□□□T1□□

7.2.1 □□□□□□□□□□□□□□□

PostgreSQL□□□□□□□□□□Read Committed□□□□□□□□□□□□□□□□□□□

```
mydb=# SELECT name,setting FROM pg_settings WHERE name =
'default_transaction_isolation';
                  name              |     setting
-------------------------------+------------------
    default_transaction_isolation | repeatable read
(1 row)
□□
mydb=# SELECT
current_setting('default_transaction_isolation');
    current_setting
-----------------
    repeatable read
(1 row)
```

7.2.2 □□□□□□□□□□□

□□1□□□□□postgresql.conf□□□□default_transaction_isolation□□□□□□□□□□□□□□□□□reload□□□□□□□

□□2□□□ALTER SYSTEM□□□□□□□□□□□□□

```
mydb=# ALTER SYSTEM SET default_transaction_isolation TO
'REPEATABLE READ';
ALTER SYSTEM
mydb=# SELECT pg_reload_conf();
    pg_reload_conf
----------------
    t
(1 row)
mydb=# SELECT current_setting('transaction_isolation');
    current_setting
-----------------
    repeatable read
(1 row)
```

7.2.3 [illegible]

[illegible]

```
mydb=# SHOW transaction_isolation ;
    transaction_isolation
-----------------------
    read committed
(1 row)
```

[illegible]

```
mydb=# SELECT current_setting('transaction_isolation');
    current_setting
-----------------
    read committed
(1 row)
```

7.2.4 [illegible]

[illegible]

```
mydb=# SET SESSION CHARACTERISTICS AS TRANSACTION ISOLATION
LEVEL READ UNCOMMITTED;
SET
mydb=# SHOW transaction_isolation ;
    transaction_isolation
-----------------------
    read uncommitted
(1 row)
```

7.2.5 [illegible]

[illegible]

```
mydb=# START TRANSACTION ISOLATION LEVEL READ UNCOMMITTED;
START TRANSACTION
...
...
...
mydb=# END;
COMMIT
```

[illegible]

```
mydb=# BEGIN ISOLATION LEVEL READ UNCOMMITTED READ WRITE;
...
...
...
mydb=# END/COMMIT/ROLLBACK;
```

START TRANSACTION[illegible]BEGIN[illegible]

7.3 PostgreSQL[illegible]

[illegible]Lock-Based Concurrency Control[illegible]Multi-Version Concurrency Control[illegible]“[illegible]”[illegible]Optimistic Concurrency Control[illegible]“OCC”[illegible]“[illegible]”[illegible]Pessimistic Concurrency Control[illegible]“PCC”[illegible]

7.3.1 [illegible]

[illegible]“[illegible]”[illegible]Exclusive locks[illegible]X[illegible]Share locks[illegible]S[illegible]

[illegible]

[illegible]

[illegible]（Granularity）[illegible]（Locking Protocol）[illegible]

[illegible]PostgreSQL[illegible]Advisory Lock[illegible]

7.3.2 [illegible]

[illegible]MVCC[illegible]MVCC[illegible]

□□□□□□□□MVCC□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□Oracle□
PostgreSQL□MySQL□Innodb□□□□MVCC□□□□
□□□□MVCC□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

MVCC□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□READ COMMITTED□□□□□□□□□□□□□
SQL□□□□□□SERIALIZABLE□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□

PostgreSQL□□□□□□□□□□□□□□□□□int32□
□□□□□□□□□□ID□□□xid□□□□□□□□□□□□□□□□□□
□□□□□id□□□□□□□□□id□□□□□□□□□□□
PostgreSQL□□□□□□□□□□□□□□□□□□□□□
PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□
□□□□□□□□□□□□□□4□□□□□□□□□□□□□□□□□
xmin□xmax□cmin□cmax□□□cmin□cmax□□
□□□□□□□□□□□□□□□□□□□□□□□□xmin□xmax□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□SQL
□□□□□□□□□□□□□□□

```
mydb=# SELECT xmin,xmax,cmin,cmax,id,ival FROM tbl_mvcc
WHERE id = 1;
    xmin | xmax | cmin | cmax | id | ival
---------+------+------+------+----+------
    1930 |    0 |    0 |    0 |  1 |    1
(1 row)
```

□□xmin□□□□□□□□□□□□□xid□xmax□□□□□□□□□xid□PostgreSQL□□□□□□□□□xmin□xmax□□□□□□□□□□□□□□

1.□□xmin□□□□□□□□

□□□□□□□□□PostgreSQL□□□□□□□□□□□□xid□□□xmin□□□□xmin□□□□□□□□□□□□□□□□□□□□□□□□□□□

1□□□

```
mydb=# BEGIN;
mydb=# SELECT txid_current();
    txid_current
--------------
         1937
(1 row)
mydb=# INSERT INTO tbl_mvcc(id,ival) VALUES(7,7);
INSERT 0 1
```

```
mydb=# SELECT xmin,xmax,cmin,cmax,id,ival FROM tbl_mvcc
WHERE id = 7;
    xmin | xmax | cmin | cmax | id | ival
---------+------+------+------+----+------
    1937 |    0 |    0 |    0 |  7 |    7
(1 row)
```

□□SELECT txid_current□□□□□□□□□□□□□□xid□1937□□□□□id□□7□□□□□□□□□□□□□□□□□□□xmin□□□□1937□□□□□□□□□□□□□□xid□

□□□□□□□□□□□□□□

```
mydb=# BEGIN;
BEGIN
mydb=# SELECT txid_current();
    txid_current
--------------
         1938
(1 row)

mydb=# SELECT * FROM tbl_mvcc WHERE id = 7;
    id | ival
-------+------
(0 rows)

mydb=# END;
COMMIT
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

2□□□□□□□□□□□□□□xid□□□□□□□Repeatable Read□Serializable□□□□□□□□□□□□xid□□□□□□□□□xid□□□□□□□xmin□□□□□□□□□xmin□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mydb=# BEGIN TRANSACTION ISOLATION LEVEL REPEATABLE READ;
BEGIN
mydb=# SELECT txid_current();
    txid_current
--------------
         1939
(1 row)
```

□□□□□□□□□□□□□□□□□xid□1939□□□□□□□□□□□□□□□

```
mydb=# BEGIN;
BEGIN
mydb=# SELECT txid_current();
    txid_current
--------------
        1940
(1 row)

mydb=# INSERT INTO tbl_mvcc (id,ival) VALUES (7,7);
INSERT 0 1
mydb=# SELECT xmin,xmax,cmin,cmax,id,ival FROM tbl_mvcc
WHERE id = 7;
    xmin | xmax | cmin | cmax | id | ival
---------+------+------+------+----+------
    1940 |    0 |    0 |    0 |  7 |    7
(1 row)
mydb=# COMMIT;
COMMIT
```

[illegible]xid[illegible]1940[illegible]xmin[illegible]xid[illegible]

```
mydb=# SELECT xmin,xmax,cmin,cmax,id,ival FROM tbl_mvcc
WHERE id = 7;
    xmin | xmax | cmin | cmax | id | ival
---------+------+------+------+----+------
(0 rows)
mydb=# END;
COMMIT
```

[illegible]XID[illegible]1939[illegible]xmin[illegible]1940[illegible]xmin[illegible]id[illegible]7[illegible]

2.[illegible]xmax[illegible]

[illegible]xmax[illegible]1[illegible]xmax[illegible]2[illegible]xid[illegible]3[illegible]COMMIT[illegible]ROLLBACK[illegible]xid[illegible]4[illegible]xid[illegible]

7.3.3 [illegible]pageinspect[illegible]MVCC

□□PostgreSQL□□□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□□pageinspect□□□□□□□□□□□□□□□□□pageinspect□□

```
mydb=# CREATE EXTENSION pageinspect;
CREATE EXTENSION
mydb=# \dx+ pageinspect
        Objects in extension "pageinspect"
            Object description
-----------------------------------------------------------------
    ...
    function get_raw_page(text,integer)
    ...
    function heap_page_items(bytea)
    ...
(19 rows)
```

□□□□□□□□□□□□get_raw_page get_raw_page□relname text□fork text□blkno int□□□□□□□□get_raw_page□relname text□blkno int□□□□□□relation□□□□□□□□□□relname□relation name□□□fork□□□main□vm□fsm□init□□□□□fork□□□□main□main□□□□□□□□□□□vm□□□□□□□□□□□fsm□free space map□□□□□init□□□□□□□get_raw_page□□□bytea□□□□□□□□□□□heap_page_items heap_page_items□□□

□MVCC□□□□□□□□□□□□□□□□□□□
get_raw_page□□□□□□□□□□□□□□□□
heap_page_items□

□□□□□□□□□□□□□□□□□PostgreSQL□MVCC□□□□□□□□□□□□□□□□□□□BRUCE MOMJIAN□□□□□□□□□□□□□□□http://momjian.us/ □

```
DROP VIEW IF EXISTS v_pageinspect;
CREATE VIEW v_pageinspect AS
SELECT  '(0,' || lp || ')' AS ctid,
        CASE lp_flags
            WHEN 0 THEN 'Unused'
            WHEN 1 THEN 'Normal'
            WHEN 2 THEN 'Redirect to ' || lp_off
            WHEN 3 THEN 'Dead'
        END,
        t_xmin::text::int8 AS xmin,
        t_xmax::text::int8 AS xmax,
        t_ctid
    FROM heap_page_items(get_raw_page('tbl_mvcc', 0))
    ORDER BY lp;
```

□□□□□□□□□□□□INSERT□□□□□□□□INSERT□□□□xmin□□□□□□□□□□xid□xmax□□□NULL□□□□□□

```
mydb=# BEGIN;
mydb=# SELECT txid_current();
    txid_current
--------------
        565
(1 row)
```

```
-- □□□□□id□565
mydb=# INSERT INTO tbl_mvcc (ival) VALUES (1);
mydb=# SELECT * FROM v_pageinspect;
    ctid  |  case  | xmin | xmax | t_ctid
----------+--------+------+------+--------
    (0,1) | Normal |  565 |    0 | (0,1)
(1 row)
mydb=# END;
-- □□□INSERT□□□□□□□xmin□□□□□□□□□□xid□565□xmax□□□NULL□
```

□DELETE□□□□□xmax□□□□□□□□□□xid□□□□□□

```
mydb=# BEGIN;
BEGIN
mydb=# SELECT txid_current();
    txid_current
--------------
        561
(1 row)

mydb=# DELETE FROM tbl_mvcc WHERE id = 1;
DELETE 1
mydb=# SELECT * FROM v_pageinspect;
    ctid  |  case  | xmin | xmax | t_ctid
----------+--------+------+------+--------
    (0,1) | Normal |  565 |  566 | (0,1)
(1 row)

mydb=# END;
COMMIT
```

□UPDATE□□□□□□□□□□□□□□□DELETE□□□□□□□□INSERT□□□□□□

```
mydb=# INSERT INTO tbl_mvcc (ival) VALUES (2); -- □□□□□□□□□□
□UPDATE□□□
mydb=# BEGIN;
mydb=# SELECT txid_current();
    txid_current
--------------
        639
(1 row)
-- □□□□□id□639
mydb=# SELECT * FROM tbl_mvcc;
    id | ival
-------+------
     2 |    2
(1 row)
-- □□tbl_mvcc□□□□□□□
mydb=# SELECT * FROM v_pageinspect;
    ctid  |  case  | xmin | xmax | t_ctid
----------+--------+------+------+--------
    (0,1) | Normal |  567 |    0 | (0,1)
(1 row)
-- □□pageinspect□□page□□□□□□□□□xmin□□□□□□xid□□□□□□□□□□□□□□
   xid□567□
mydb=# UPDATE tbl_mvcc SET ival = 20 WHERE id = 2;
-- □□□□□□
mydb=# SELECT * FROM v_pageinspect;
    ctid  |  case  | xmin | xmax | t_ctid
----------+--------+------+------+--------
    (0,1) | Normal |  567 |  639 | (0,2)
    (0,2) | Normal |  639 |    0 | (0,2)
(2 rows)
mydb=# END;
```

□□pageinspect□□page□□□□□□□□UPDATE□□□□□DELETE□□□□□□□INSERT□□□□□□□□□□□□□□□□□□□□□xid□567□□□□□ctid□□0□1□□□□□□□xmin□567□xmax□□□□□□□xid□639□□□□□□page□□□□□ctid□□0□2□□□□□□□xmin□□□□□□□xid□639□□□□□□□□□□□□□□□□□□□

□□□UPDATE□□□□□□□□□□□□□UPDATE□□□□□□
□□□□□□□

7.3.4 □□pg_repack□□□□□□□

□□PostgreSQL□MVCC□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□PostgreSQL□□□□□□□□□□□□MVCC□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL
□MVCC□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□

□□VACUUM□□□□autovacuum□□□□□□□□
□□□□□□□□□□□VACUUM□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□VACUUM FULL□□□□□□
□□□□□□□□□□□□□□□□□□□□□

pg_repack□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□COPY□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□yum
□□□□□□□□□yum□□□□□□□□□□pg_repack□□□□

□□

```
[root@pghost1 ~]# yum install -y pg_repack10
```

□□□□□□□□□pg_repack□□□□□□□□

```
mydb=# CREATE EXTENSION pg_repack;
```

□□□□□□□□pg_repack□tbl_mvcc□□□□□□□□□□□

```
[pos tgres@pghost1 ~]# /usr/pgsql-10/bin/pg_repack -t
tbl_mvcc -j 2 -D -k -h pghost1-U postgres -d mydb
```

□□□PostgreSQL□□□□□□□□□□□□□□□□□□□MVCC□□□□□□□□□□□□□□□□□□□

7.3.5 □□□□□DDL

PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□DDL□□□□□DDL□□□□□□□□□□□□□□□□TRUNCATE□□□□□□□□□□□□□□□□□

```
mydb=# DROP TABLE IF EXISTS tbl_test;
NOTICE:  table "tbl_test" does not exist, skipping
DROP TABLE
mydb=# BEGIN;
BEGIN
mydb=# CREATE TABLE tbl_test (ival int);
CREATE TABLE
mydb=# INSERT INTO tbl_test VALUES (1);
INSERT 0 1
mydb=# ROLLBACK;
ROLLBACK
mydb=# SELECT * FROM tbl_test;
ERROR:  relation "tbl_test" does not exist
```

[illegible]TRUNCATE[illegible]

```
mydb=# SELECT COUNT(*) FROM tbl_mvcc;
 count
-------
 9
(1 row)
mydb=# BEGIN;
BEGIN
mydb=# TRUNCATE tbl_mvcc ;
TRUNCATE TABLE
mydb=# ROLLBACK;
ROLLBACK
mydb=# SELECT COUNT(*) FROM tbl_mvcc;
 count
-------
 9
(1 row)
```

[illegible]TRUNCATE[illegible]

7.4 □□□□

□□PostgreSQL□□PostgreSQL□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□pageinspect□□□□□□□□□□MVCC□□□□□□□□□□□□PostgreSQL□□□□□□□□DDL□

[illegible]8[illegible] [illegible]

[illegible]Oracle[illegible]PostgreSQL[illegible]10[illegible]PostgreSQL[illegible]PostgreSQL10[illegible]10[illegible]

8.1 [illegible]

[illegible]

·[illegible]

·[illegible]DROP TABLE[illegible]DELETE[illegible]

·[illegible]

[illegible]

8.2 [illegible]

[illegible]

8.2.1 [illegible]

PostgreSQL[illegible]

[illegible]tbl_log[illegible]

```
mydb=> CREATE TABLE tbl_log(id int4,create_date date,log_type
text);
CREATE TABLE
```

[illegible]tbl_log_sql[illegible]SQL[illegible]

```
mydb=> CREATE TABLE tbl_log_sql(sql text) INHERITS(tbl_log);
CREATE TABLE
```

[illegible]INHERITS[illegible]tbl_log[illegible]tbl_log_sql[illegible]tbl_log[illegible]sql[illegible]

□□□□□tbl_log□□□tbl_log_sql□□□□□□□□□

```
mydb=> \d tbl_log_sql
            Table "pguser.tbl_log_sql"
        Column  |  Type   | Collation | Nullable | Default
---------------+---------+-----------+----------+---------
    id          | integer |           |          |
    create_date | date    |           |          |
    log_type    | text    |           |          |
    sql         | text    |           |          |
Inherits: tbl_log
```

□□□□□tbl_log_sql□□□□□□□□□□□□□□
tbl_log□□□□□□□□sql□□□□□□□□□Inherits□
tbl_log□□□□□□□□tbl_log□

□□□□□□□□□□□□□□□□
□□□

```
mydb=> INSERT INTO tbl_log VALUES (1,'2017-08-26',null);
INSERT 0 1
mydb=> INSERT INTO tbl_log_sql VALUES(2,'2017-08-
27',null,'select 2');
INSERT 0 1
```

□□□□□□□□tbl_log□□□□□□□□□□□□□

```
mydb=> SELECT * FROM tbl_log;
    id | create_date | log_type
-------+-------------+----------
     1 | 2017-08-26  |
```

```
     2 | 2017-08-27  |
(2 rows)
```

[illegible]SQL[illegible]OID[illegible]

```
mydb=> SELECT tableoid, * FROM tbl_log;
    tableoid | id | create_date | log_type
-------------+----+-------------+----------
       16854 |  1 | 2017-08-26  |
       16860 |  2 | 2017-08-27  |
(2 rows)
```

tableoid[illegible]OID[illegible]pg_class[illegible]

```
mydb=> SELECT p.relname,c.*
       FROM tbl_log c, pg_class p
       WHERE c.tableoid = p.oid;
        relname | id | create_date | log_type
----------------+----+-------------+----------
    tbl_log     |  1 | 2017-08-26  |
    tbl_log_sql |  2 | 2017-08-27  |
(2 rows)
```

[illegible]ONLY[illegible]

```
mydb=> SELECT * FROM ONLY tbl_log;
    id | create_date | log_type
-------+-------------+----------
```

```
      1 | 2017-08-26  |
(1 row)
```

[illegible]UPDATE[illegible]DELETE[illegible]SELECT[illegible]ONLY[illegible]DML[illegible]

```
mydb=> DELETE FROM tbl_log;
DELETE 2
mydb=> SELECT count(*) FROM tbl_log;
 count
-------
 0
(1 row)
```

[illegible]

[illegible] [illegible]UPDATE[illegible]DELETE[illegible]DML[illegible]

8.2.2 [illegible]

[illegible]

[illegible]1 [illegible]

[illegible]2 [illegible]INHERITS[illegible]

[illegible]

[illegible]3 [illegible]

[illegible]

[illegible]4 [illegible]

[illegible]

[illegible]5 [illegible]INSERT[illegible]DELETE[illegible]UPDATE[illegible]SQL[illegible]DML[illegible]

[illegible]6 [illegible]constraint_exclusion[illegible]off[illegible]SQL[illegible]

[illegible]

[illegible]

```
CREATE TABLE log_ins(id serial,
user_id int4,
create_time timestamp(0) without time zone);
```

[illegible]13[illegible]

```
CREATE TABLE log_ins_history(CHECK ( create_time < '2017-01-
01' )) INHERITS(log_ins);
CREATE TABLE log_ins_201701(CHECK ( create_time >= '2017-01-
01' and create_time < '2017-02-01')) INHERITS(log_ins);
CREATE TABLE log_ins_201702(CHECK ( create_time >= '2017-02-
01' and create_time < '2017-03-01')) INHERITS(log_ins);
...
CREATE TABLE log_ins_201712(CHECK ( create_time >= '2017-12-
01' and create_time < '2018-01-01')) INHERITS(log_ins);
```

```
CREATE INDEX idx_his_ctime ON log_ins_history USING btree
(create_time);
CREATE INDEX idx_log_ins_201701_ctime ON log_ins_201701 USING
btree (create_time);
CREATE INDEX idx_log_ins_201702_ctime ON log_ins_201702 USING
btree (create_time);
...
CREATE INDEX idx_log_ins_201712_ctime ON log_ins_201712 USING
btree (create_time);
```

```
CREATE OR REPLACE FUNCTION log_ins_insert_trigger()
    RETURNS trigger
    LANGUAGE plpgsql
AS $function$
BEGIN
    IF    ( NEW. create_time < '2017-01-01' ) THEN
        INSERT INTO log_ins_history VALUES (NEW.*);
    ELSIF ( NEW.create_time>='2017-01-01' and
NEW.create_time<'2017-02-01' ) THEN
        INSERT INTO log_ins_201701 VALUES (NEW.*);
    ELSIF ( NEW.create_time>='2017-02-01' and
```

```
NEW.create_time<'2017-03-01' ) THEN
        INSERT INTO log_ins_201702 VALUES (NEW.*);
    ...
    ELSIF ( NEW.create_time>='2017-12-01' and
NEW.create_time<'2018-01-01' ) THEN
        INSERT INTO log_ins_201712 VALUES (NEW.*);
    ELSE
        RAISE EXCEPTION 'create_time out of range.  Fix the
log_ins_insert_trigger() function!';
    END IF;
    RETURN NULL;
END;
$function$;
```

[illegible]new.*[illegible]

```
CREATE TRIGGER insert_log_ins_trigger BEFORE INSERT ON
log_ins FOR EACH ROW EXECUTE PROCEDURE
log_ins_insert_trigger();
```

[illegible]log_ins[illegible]log_ins_insert_trigger[illegible]DELETE[illegible]UPDATE[illegible]INSERT[illegible]

[illegible] [illegible]user_id[illegible]user_id[illegible]

8.2.3 □□□□□

□□□log_ins□□□□□□□□□□□□□□□□□□□□□□□□□□

```
INSERT INTO log_ins(user_id,create_time)
SELECT round(100000000*random()),generate_series('2016-12-
01'::date,
    '2017-12-01'::date, '1 minute');
```

□□□□round□100000000*random□□□□□□□8□□□□generate_series□□□□□□□□□□□□□□□□

```
mydb=> SELECT * FROM log_ins LIMIT 2;
        id | user_id  |     create_time
-----------+----------+---------------------
    570242 | 24040985 | 2016-12-01 00:00:00
    570243 | 10814368 | 2016-12-01 00:01:00
(2 rows)
```

□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT count(*) FROM ONLY log_ins;
  count
-------
  0
(1 row)

mydb=> SELECT count(*) FROM log_ins;
  count
--------
  525601
(1 row)
```

[illegible]

```
mydb=> SELECT min(create_time),max(create_time) FROM
log_ins_201701;
        min              |         max
------------------------+---------------------
    2017-01-01 00:00:00 | 2017-01-31 23:59:00
(1 row)
```

[illegible]

```
mydb=> \dt+ log_ins*
                              List of relations
    Schema |      Name       | Type  | Owner  |    Size    |
Description
-----------+-----------------+-------+--------+------------+-
------------
    pguser | log_ins         | table | pguser | 0 bytes    |
    pguser | log_ins_201701  | table | pguser | 1960 kB    |
    pguser | log_ins_201702  | table | pguser | 1768 kB    |
...
pguser | log_ins_201712  | table | pguser | 8192 bytes |
pguser | log_ins_history | table | pguser | 1960 kB    |
(14 rows)
```

[illegible]

8.2.4 [illegible]

[illegible]2017-01-01[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT * FROM log_ins WHERE
create_time > '2017-01-01' AND create_time < '2017-01-02';
QUERY PLAN
-------------------------------------------------------------
--------------------
    Append  (cost=0.00..45.97 rows=1435 width=16) (actual
time=0.025..0.425 rows=1439 loops=1)
        ->  Seq Scan on log_ins  (cost=0.00..0.00 rows=1
width=16) (actual time=0.004.. 0.004 rows=0 loops=1)
            Filter: ((create_time > '2017-01-01
00:00:00'::timestamp without time zone) AND (create_time <
'2017-01-02 00:00:00'::times
tamp without time zone))
        ->  Index Scan USING idx_log_ins_201701_ctime on
log_ins_201701  (cost=0.29.. 45.97 rows=1434 width=16)
(actual time=0.020..0.285
rows=1439 loops=1)
            Index Cond: ((create_time > '2017-01-01
00:00:00'::timestamp without time zone) AND (create_time <
'2017-01-02 00:00:00'::t
imestamp without time zone))
    Planning time: 0.581 ms
    Execution time: 0.515 ms
(7 rows)
```

[illegible]log_ins_201701[illegible]0.515[illegible]log_ins_201701[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT * FROM log_ins_201701 WHERE
create_time > '2017-01-01' AND create_time < '2017-01-02';
QUERY PLAN
-------------------------------------------------------------
---------
    Index Scan USING idx_log_ins_201701_ctime on
log_ins_201701  (cost=0.29..45.97 rows=1434 width=16) (actual
time=0.017..0.254 rows=1
439 loops=1)
        Index Cond: ((create_time > '2017-01-01
```

```
00:00:00'::timestamp without time zone) AND (create_time <
'2017-01-02 00:00:00'::timestamp
            without time zone))
    Planning time: 0.142 ms
    Execution time: 0.337 ms
(4 rows)
```

[illegible]0.337[illegible]

log_ins[illegible]

8.2.5 constraint_exclusion[illegible]

constraint_exclusion[illegible]

·on[illegible]

·off[illegible]

·partition[illegible]UNION ALL[illegible]

[illegible]on[illegible]partition[illegible]

□□□□□□□□□constraint_exclusion□□□off□□□□□□□□□□□

```
mydb=> SET constraint_exclusion =off;
SET
```

□□□□□□□□□□□□□

```
mydb=> EXPLAIN ANALYZE SELECT * FROM log_ins WHERE
create_time > '2017-01-01' AND create_time < '2017-01-02';
QUERY PLAN
-------------------------------------------------------------
--------------------
    Append  (cost=0.00..94.40 rows=1447 width=16) (actual
time=0.029..0.534 rows=1439 loops=1)
            ->  Seq Scan on log_ins  (cost=0.00..0.00 rows=1
width=16) (actual time=0.005..0.005 rows=0 loops=1)
         Filter: ((create_time > '2017-01-01
00:00:00'::timestamp without time zone) AND (create_time <
'2017-01-02 00:00:00'::timestamp without time zone))
            ->  Index Scan USING idx_his_ctime on
log_ins_history  (cost=0.29..4.31 rows=1 width=16) (actual
time=0.008..0.008 rows=0 loops=1)
                Index Cond: ((create_time > '2017-01-01
00:00:00'::timestamp without time zone) AND (create_time <
'2017-01-02 00:00:00'::
                     timestamp without time zone))
            ->  Index Scan USING idx_log_ins_201701_ctime on
log_ins_201701  (cost=0.29..45.97 rows=1434 width=16) (actual
time=0.016..0.293
                  rows=1439 loops=1)
                Index Cond: ((create_time > '2017-01-01
00:00:00'::timestamp without time zone) AND (create_time <
'2017-01-02 00:00:00'::
                    timestamp without time zone))=0 loops=1)
        ...
            ->  Seq Scan on log_ins_201712  (cost=0.00..1.01
rows=1 width=16) (actual time=0.011..0.011 rows=0 loops=1)
                Filter: ((create_time > '2017-01-01
```

```
00:00:00'::timestamp without time zone) AND (create_time <
'2017-01-02 00:00:00'::
                    timestamp without time zone))
                Rows Removed by Filter: 1
    Planning time: 1.344 ms
    Execution time: 0.685 ms
(32 rows)
```

[illegible]0.685[illegible]partition[illegible]on[illegible]

8.2.6 [illegible]

[illegible]log_ins[illegible]SQL[illegible]

```
CREATE TABLE log_ins_201801(CHECK ( create_time >= '2018-01-
01' and create_time < '2018-02-01')) INHERITS(log_ins);
...
```

[illegible]

[illegible]

```
CREATE INDEX idx_log_ins_201801_ctime ON log_ins_201801 USING
btree
(create_time);
...
```

[illegible]log_ins_insert_trigger[illegible]\d+log_ins[illegible]log_ins[illegible]

[illegible]

```
--[illegible]
CREATE TABLE log_ins_201802(LIKE log_ins INCLUDING ALL );

--[illegible]
ALTER TABLE log_ins_201802 ADD CONSTRAINT
log_ins_201802_create_time_check
 CHECK ( create_time >= '2018-02-01' AND create_time < '2018-
03-01');

--[illegible]log_ins_insert_trigger()
[illegible]

--[illegible]log_ins_201802[illegible]log_ins
ALTER TABLE log_ins_201802 INHERIT log_ins;
```

[illegible]

8.2.7 [illegible]

[illegible]DELETE[illegible]

```
DROP TABLE log_ins_201802
```

[illegible]

```
mydb=> ALTER TABLE log_ins_201802 NO INHERIT log_ins;
ALTER TABLE
```

[illegible]log_ins_201802[illegible] log_ins[illegible]log_ins_201802[illegible]

8.2.8 [illegible]

[illegible]\d[illegible]

```
mydb=> \d log_ins
Table "pguser.log_ins"
```

```
   Column    |   Type   | Collation | Nullable |
Default
------------+---------+----------+---------+------------
-------------------
 id          | integer  | not null  |
nextval('log_ins_id_seq'::regclass)
 user_id     | integer  |           |          |
 create_time | timestamp(0) without time zone  |           |
|
Triggers:
    insert_log_ins_trigger BEFORE INSERT ON log_ins FOR EACH
ROW EXECUTE PROCEDURE log_ins_insert_trigger()
Number of child tables: 14 (Use \d+ to list them.)
```

□□□□□□□□log_ins□14□□□□□□□□□□□□□□□□□□□log_ins_insert_trigger□□□□□□□□□□□□□□□\d+log_ins□□□□□□

□□□□□□□□□□□□□□□□□SQL□□□□□□□□

```
mydb=> SELECT
    nmsp_parent.nspname AS parent_schema ,
    parent.relname AS parent ,
    nmsp_child.nspname AS child_schema ,
    child.relname AS child_schema
FROM
    pg_inherits JOIN pg_class parent
        ON pg_inherits.inhparent = parent.oid JOIN pg_class
child
        ON pg_inherits.inhrelid = child.oid JOIN pg_namespace
nmsp_parent
        ON nmsp_parent.oid = parent.relnamespace JOIN
pg_namespace nmsp_child
        ON nmsp_child.oid = child.relnamespace
WHERE
    parent.relname = 'log_ins';
    parent_schema | parent  | child_schema |  child_schema
----------------+--------+-------------+----------------
    pguser        | log_ins | pguser       | log_ins_history
```

```
    pguser       | log_ins | pguser       | log_ins_201701
    pguser       | log_ins | pguser       | log_ins_201702
    ...
    pguser       | log_ins | pguser       | log_ins_201801
(14 rows)
```

pg_inherits[illegible]SQL[illegible]

```
mydb=> SELECT
    nspname ,
    relname ,
    count(*) AS partition_num
FROM
    pg_class c ,
    pg_namespace n ,
    pg_inherits i
WHERE
    c.oid = i.inhparent
    AND c.relnamespace = n.oid
    AND c.relhassubclass
    AND c.relkind in ('r','p')
GROUP BY 1,2 ORDER BY partition_num DESC;
    nspname | relname | partition_num
------------+---------+---------------
    pguser  | log_ins |            14
    pguser  | tbl_log |             1
(2 rows)
```

[illegible]log_ins[illegible]14[illegible]tbl_log[illegible]

8.2.9 [illegible]

[illegible]create_time[illegible]log_ins[illegible]user_id[illegible]

[illegible]log[illegible]log_ins[illegible]

```
CREATE TABLE log(id
serial,user_id int4,
create_time timestamp(0) without time zone);

INSERT INTO log(user_id,create_time)
SELECT round(100000000*random()),generate_series('2016-12-
01'::date,
'2017-12-01'::date, '1 minute');
```

[illegible]

```
mydb=> SELECT count(*) FROM log_ins;
 count
--------
 525601
(1 row)

mydb=> SELECT count(*) FROM log;
 count
--------
 525601
(1 row)
```

[illegible]log[illegible]

```
CREATE INDEX idx_log_userid ON log USING btree(user_id);
CREATE INDEX idx_log_create_time ON log USING
btree(create_time);
```

□□□□log_ins□□□□□□□□user_id□□□□□□□□□□□

```
CREATE INDEX idx_log_ins_userid ON log_ins USING
btree(user_id);
CREATE INDEX idx_his_userid ON log_ins_history USING btree
(user_id);
CREATE INDEX idx_log_ins_201701_userid ON log_ins_201701
USING btree (user_id);
CREATE INDEX idx_log_ins_201702_userid ON log_ins_201702
USING btree (user_id);
...
CREATE INDEX idx_log_ins_201801_userid ON log_ins_201801
USING btree (user_id);
```

□□□□□user_id□□□□□□□□□□log_ins□□□□□□□□□□□□□user_id□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□

```
--□□□□□user_id□□
SELECT * FROM log WHERE user_id=?;
SELECT * FROM log_ins WHERE user_id=?;
```

□□□□□□□□log□log_ins□□□□user_id□□□□□□

```
mydb=> SELECT a.* FROM log a ,log_ins b WHERE
a.user_id=b.user_id LIMIT 1;
    id    | user_id  |     create_time
----------+----------+---------------------
    67286 | 51751630 | 2017-01-16 17:25:00
(1 row)
```

□□user_id=51751630□□□□□□□log□□□□□□□□□

```
mydb=> EXPLAIN SELECT * FROM log WHERE user_id=51751630;
                                QUERY PLAN
-------------------------------------------------------------
--------------
    Index Scan using idx_log_userid on log  (cost=0.42..4.44
rows=1 width=16)
        Index Cond: (user_id = 51751630)
(2 rows)
```

□□□□□□□□□□□□□□□□user_id□□□□□□□□□□□□□□□□□□□□□

```
mydb=> EXPLAIN SELECT * FROM log_ins WHERE user_id=51751630;
                                    QUERY PLAN
-------------------------------------------------------------
-------------------
 Append  (cost=0.00..63.18 rows=23 width=16)
    ->  Seq Scan on log_ins  (cost=0.00..0.00 rows=1
width=16)
        Filter: (user_id = 88258037)
    ->  Index Scan USING idx_his_userid on log_ins_history
          (cost=0.29..4.31 rows=1 width=16)
        Index Cond: (user_id = 88258037)
    ->  Index Scan USING idx_log_ins_201701_userid on
log_ins_201701
          (cost=0.29.. 4.31 rows=1 width=16)
        Index Cond: (user_id = 88258037)
```

```
    ->  Index Scan USING idx_log_ins_201702_userid on
log_ins_201702  (cost=0.29.. 4.31 rows=1 width=16)
        Index Cond: (user_id = 88258037)
    ...
    ->  Bitmap Heap Scan on log_ins_201801  (cost=2.22..10.48
rows=9 width=16)
        Recheck Cond: (user_id = 88258037)
            ->  Bitmap Index Scan on
idx_log_ins_201801_userid  (cost=0.00..2.22 rows=9 width=0)
                Index Cond: (user_id = 88258037)
(33 rows)
```

□□□□□□□□□□□□□□□□□□□□□□log□□□□SQL□□□□□□□□□0.050□□□log_ins□□□□SQL□□□□□□□□□0.184□□

create_time□□□□log_ins□□□□□□□□SQL□□□□

```
--□□□□□create_time□□;
SELECT * FROM log WHERE create_time > '2017-01-01' AND
create_time < '2017-01-02';
SELECT * FROM log_ins WHERE create_time > '2017-01-01' AND
create_time < '2017-01-02';
```

□log□□□□SQL□□□□□□□□□0.339□□□log_ins□□□□SQL□□□□□□□□□□0.503□□□8-1□□□□□□□□□□□□

□8-1　□□□□□□□□□

查询场景	普通表 log 执行时间	分区表：查询 log_ins 父表执行时间	分区表：查询 log_ins 子表执行时间
根据非分区键 user_id 查询	0.05 毫秒	0.184 毫秒	不支持
根据分区键 create_time 范围查询	0.339 毫秒	0.503 毫秒	0.325 毫秒

[illegible]user_id[illegible]2.68[illegible]create_time[illegible]0.4[illegible]create_time[illegible]user_id[illegible]

1[illegible]

2[illegible]

[illegible]90%[illegible]SQL[illegible]

8.2.10 [illegible]

[illegible]

·[illegible]

·[illegible]

·[illegible]

·UPDATE[illegible]

·[illegible]

8.3 □□□□□

PostgreSQL10□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□INSERT□DELETE□UPDATE□□□□□□□□DML□□□PostgreSQL10□□□□□□□□□□□□□□□□□

8.3.1 □□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□

```
CREATE TABLE table_name ( ... )
    [ PARTITION BY { RANGE | LIST } ( { column_name | (
expression ) }
```

□□□□□□□□□□□□□□□□□□□□□RANGE□□□□□LIST□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□

```
CREATE TABLE table_name
    PARTITION OF parent_table [ (
    ) ] FOR VALUES partition_bound_spec
```

□□□□□□□□□□□□□□□□□□□
partition_bound_spec□□□□□□□
partition_bound_spec□□□□□□□□□□□□□□
□□□□partition_bound_spec□□□□□□□□□
□□

PostgreSQL10□□□□□□□□□□□□□□

1□□□□□□□□□□□□□

2□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□

3□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□log_par□□□□□

```
CREATE TABLE log_par (
    id serial,
    user_id int4,
    create_time timestamp(0) without time zone
) PARTITION BY RANGE(create_time);
```

□log_par□□□□□□□□□□□□□
create_time□□

[illegible]

```
CREATE TABLE log_par_his PARTITION OF log_par FOR VALUES
FROM (UNBOUNDED)
TO ('2017-01-01');
CREATE TABLE log_par_201701 PARTITION OF log_par FOR VALUES
FROM ('2017-01-01') TO ('2017-02-01');
CREATE TABLE log_par_201702 PARTITION OF log_par FOR VALUES
FROM ('2017-02-01') TO ('2017-03-01');
...
CREATE TABLE log_par_201712 PARTITION OF log_par FOR VALUES
FROM ('2017-12-01') TO ('2018-01-01');
```

[illegible]

[illegible]

```
CREATE INDEX idx_log_par_his_ctime ON log_par_his USING
btree(create_time);
CREATE INDEX idx_log_par_201701_ctime ON log_par_201701
USING btree(create_time);
CREATE INDEX idx_log_par_201702_ctime ON log_par_201702
USING btree(create_time);
...
CREATE INDEX idx_log_par_201712_ctime ON log_par_201712
USING btree(create_time);
```

[illegible]

8.3.2 [illegible]

□□□□□□□□□□□□□□

```
INSERT INTO log_par(user_id,create_time)
SELECT round(100000000*random()),generate_series('2016-12-
01'::date,
    '2017-12-01'::date, '1 minute');
```

□□□□□□□□□□□

```
mydb=> SELECT count(*) FROM log_par;
 count
--------
 525601
(1 row)

mydb=> SELECT count(*) FROM ONLY log_par;
 count
-------
 0
(1 row)
```

□□□□□□□□□□□□log_par□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> \dt+ log_par*
                          List of relations
    Schema |      Name      | Type  | Owner  |    Size    |
Description
-----------+----------------+-------+--------+------------+-
------------
    pguser | log_par        | table | pguser | 0 bytes    |
    pguser | log_par_201701 | table | pguser | 1960 kB    |
    pguser | log_par_201702 | table | pguser | 1768 kB    |
    ...
```

```
    pguser | log_par_201712 | table | pguser | 8192 bytes |
    pguser | log_par_his    | table | pguser | 1960 kB    |
```

8.3.3 [illegible]

[illegible]log_par[illegible]

```
mydb=> SELECT
    nmsp_parent.nspname AS parent_schema ,
    parent.relname AS parent ,
    nmsp_child.nspname AS child_schema ,
    child.relname AS child_schema
FROM
    pg_inherits JOIN pg_class parent
        ON pg_inherits.inhparent = parent.oid JOIN pg_class
child
        ON pg_inherits.inhrelid = child.oid JOIN
pg_namespace nmsp_parent
        ON nmsp_parent.oid = parent.relnamespace JOIN
pg_namespace nmsp_child
        ON nmsp_child.oid = child.relnamespace
WHERE
    parent.relname = 'log_par';
 parent_schema | parent  | child_schema |  child_schema
---------------+---------+--------------+----------------
 pguser        | log_par | pguser       | log_par_his
 pguser        | log_par | pguser       | log_par_201701
 pguser        | log_par | pguser       | log_par_201702
 ...
 pguser        | log_par | pguser       | log_par_201712
(13 rows)
```

[illegible]SQL[illegible]log_par[illegible]\d+log_par[illegible]log_par[illegible]

8.3.4 □□□□

□□□□□□□□□□□□□□□log_par□□□□□□□□□□
□□

```
CREATE TABLE log_par_201801 PARTITION OF log_par FOR VALUES
FROM ('2018-01-01') TO ('2018-02-01');
```

□□□□□□□□□□□□□□□

```
CREATE INDEX idx_log_par_201801_ctime ON log_par_201801
USING btree(create_time);
```

8.3.5 □□□□

□□□□□□□□□□□□□□□□□DROP□□□□□□□□□□□
□□□

```
DROP TABLE log_par_201801□
```

DROP□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> ALTER TABLE log_par DETACH PARTITION log_par_201801;
ALTER TABLE
```

[illegible]

```
mydb=> ALTER TABLE log_par ATTACH PARTITION log_par_201801
FOR VALUES FROM ('2018-01-01') TO ('2018-02-01');
ALTER TABLE
```

[illegible]

8.3.6 [illegible]

[illegible]2017-01-01[illegible]SQL[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT * FROM log_par WHERE
create_time > '2017-01-01' AND create_time < '2017-01-02';
                                QUERY PLAN
------------------------------------------------------------
--------------------
    Append  (cost=0.29..45.21 rows=1396 width=16) (actual
time=0.019..0.425 rows=1439 loops=1)
        ->  Index Scan using idx_log_par_201701_ctime on
log_par_201701
              (cost=0.29.. 45.21 rows=1396 width=16) (actual
time=0.019..0.288
rows=1439 loops=1)
            Index Cond: ((create_time > '2017-01-01
00:00:00'::timestamp without time zone) AND (create_time <
'2017-01-02 00:00:00'::timestamp without time zone))
    Planning time: 0.461 ms
    Execution time: 0.510 ms
(5 rows)
```

□□□□□□□□□□□□□□□log_par_201701□□□□□□□□□□□□□□0.510□□□

□□□□□□□□□□□log_par□□□□log□□□□□□create_time□□□□□log_par□□□□□user_id□□□

□□□□log_par□□□□□user_id□□□□□□□□□□□

```
CREATE INDEX idx_log_par_his_userid ON log_par_his using
btree (user_id);
CREATE INDEX idx_log_par_201701_userid ON log_par_201701
using btree (user_id);
CREATE INDEX idx_log_par_201702_userid ON log_par_201702
using btree (user_id);
...
CREATE INDEX idx_log_par_201712_userid ON log_par_201712
using btree (user_id);
```

□□user_id□□□□□□□□□□log_par□□□□□□□□□□□□□□□□SQL□□□□□□

```
--□□□□□□user_id□□
SELECT * FROM log WHERE user_id=?;
SELECT * FROM log_par WHERE user_id=?;
```

□□□□□□□□□□log□log_par□□□□user_id□□□□□□

```
mydb=>  SELECT a.* FROM log a ,log_par b WHERE
a.user_id=b.user_id LIMIT 1;
        id | user_id  |     create_time
-----------+----------+---------------------
    258926 | 70971018 | 2017-05-29 19:25:00
(1 row)
```

□□user_id=70971018□□□□□□□□log□□□□□□□□□□□

```
mydb=> EXPLAIN SELECT * FROM log WHERE user_id=70971018;
                                QUERY PLAN
------------------------------------------------------------
---------------
    Index Scan using idx_log_userid on log  (cost=0.42..4.44
rows=1 width=16)
        Index Cond: (user_id = 70971018)
(2 rows)
```

□□□□□□□□□□□□□□□□

□□□□□□user_id□□□□□□□□log_par□□□□□□□□□□□□□□□□□□

```
mydb=> EXPLAIN SELECT * FROM log_par WHERE user_id=70971018;
                                 QUERY PLAN
------------------------------------------------------------
---------------------
 Append  (cost=0.29..52.70 rows=13 width=16)
    ->  Index Scan using idx_log_par_his_userid on
log_par_his  (cost=0.29..4.31 rows=1 width=16)
        Index Cond: (user_id = 70971018)
    ->  Index Scan using idx_log_par_201701_userid on
log_par_201701
```

```
          (cost=0.29.. 4.31 rows=1 width=16)
        Index Cond: (user_id = 70971018)
    ->  Index Scan using idx_log_par_201702_userid on
log_par_201702
          (cost=0.29.. 4.31 rows=1 width=16)
        Index Cond: (user_id = 70971018)
    ...
    ->  Seq Scan on log_par_201712  (cost=0.00..1.01 rows=1
width=16)
        Filter: (user_id = 70971018)
(27 rows)
```

□□□□□□□□□□□□□□□□user_id□□□□□log_par□□□□□□□□□□log□□□□□□SQL□□□□□□□□□□□□□0.047□□□log_par□□□□□□SQL□□□□□□□□□□□□□0.139□□□

create_time□□□□□□log_par□□□□□□□□□□□SQL□□□□□□

```
--□□□□□□create_time□□;
SELECT * FROM log WHERE create_time > '2017-01-01' AND
create_time < '2017-01-02';
SELECT * FROM log_par WHERE create_time > '2017-01-01' AND
create_time < '2017-01-02';
```

□log□□□□□□SQL□□□□□□□□□□□□□0.340□□□□log_par□□□□□SQL□□□□□□□□□□□□□0.503□□□□8-2□□□□□□□□□□□□□□□□□

□8-2 □□□□□□□□□□□□□

查询场景	普通表 log 执行时间	分区表：查询 log_par 父表执行时间	区表：查询 log_par 子表执行时间
根据非分区键 user_id 查询	0.047 毫秒	0.139 毫秒	不支持
根据分区键 create_time 范围查询	0.340 毫秒	0.503 毫秒	0.319 毫秒

[illegible]user_id[illegible]1.95[illegible]create_time[illegible]0.47[illegible]create_time[illegible]user_id[illegible]8-1[illegible]8-2[illegible]8-3[illegible]

[illegible]8-3 [illegible]

查询场景	普通表 (log)	传统分区表 (log_ins)		内置分区表 (log_par)	
		查询父表	查询子表	查询父表	查询子表
根据非分区键 user_id 查询	表 8-1：0.05 毫秒 表 8-2：0.047 毫秒	0.184 毫秒	不支持	0.139 毫秒	不支持
根据分区键 create_time 范围查询	表 8-1：0.339 毫秒 表 8-2：0.340 毫秒	0.503 毫秒	0.325 毫秒	0.503 毫秒	0.319 毫秒

[illegible]

·[illegible]

·[illegible]

[illegible]

8.3.7 constraint_exclusion[illegible]

[illegible]constraint_exclusion[illegible]

```
mydb=> SET constraint_exclusion =off;
SET
```

[illegible]SQL[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT * FROM log_par WHERE
create_time > '2017-01-01' AND create_time < '2017-01-02';
QUERY PLAN
------------------------------------------------------------
---------------------
 Append  (cost=0.29..104.16 rows=1417 width=16) (actual
time=0.024..0.460 rows=1439 loops=1)
    ->  Index Scan using idx_log_par_his_ctime on
log_par_his  (cost=0.29..4.31 rows=1 width=16) (actual
time=0.007..0.007 rows=0 loops=1)
            Index Cond: ((create_time > '2017-01-01
00:00:00'::timestamp without time zone) AND
(create_time < '2017-01-02 00:00:00'::timestamp without time
zone))
```

```
        ->  Index Scan using idx_log_par_201701_ctime on
log_par_201701  (cost=0.29..45.21 rows=1396 width=16)
              (actual time=0.016..0.280 rows=1439 loops=1)
         Index Cond: ((create_time > '2017-01-01
00:00:00'::timestamp without time zone) AND (create_time <
'2017-01-02 00:00:00'::timestamp without time zone))
...
            ->  Bitmap Index Scan on
idx_log_par_201801_ctime  (cost=0.00..2.24 rows=9 width=0)
(actual time=0.002..0.002 rows=0 loops=1)
                Index Cond: ((create_time > '2017-01-01
00:00:00'::timestamp without time zone) AND
                    (create_time < '2017-01-02
00:00:00'::timestamp without time zone))
 Planning time: 0.792 ms
 Execution time: 0.607 ms
(34 rows)
```

[illegible]0.607[illegible]partition[illegible]on[illegible]

8.3.8 [illegible]

[illegible]UPDATE[illegible]SQL[illegible]

```
mydb=> SELECT * FROM log_par_201701 LIMIT 1;
  id   | user_id  |     create_time
-------+----------+---------------------
 44641 | 16965492 | 2017-01-01 00:00:00
(1 row)

mydb=> UPDATE log_par SET create_time='2017-02-02 01:01:01'
```

```
WHERE user_id=16965492;
ERROR:  new row for relation "log_par_201701" violates
partition constraint
DETAIL:  Failing row contains (44641, 16965492, 2017-02-02
01:01:01).
```

□□user_id□□16965492□□□□□log_par_201701□□□□□□□□□create_time□□□'2017-02-0201□01□01'□□□

```
mydb=> UPDATE log_par SET create_time='2017-01-01 01:01:01'
WHERE user_id=16965492;
UPDATE 1
```

□□

8.3.9 □□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

·□□□

·□□

·[illegible]

·[illegible]

·[illegible]UPDATE[illegible]
UPDAET[illegible]

·[illegible]

8.4 [illegible]

[illegible]PostgreSQL10[illegible]SQL[illegible]

□9□ PostgreSQL□NoSQL□□

PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□3□□□□□□□□□□□PostgreSQL□json□jsonb□□□□□□□□□□□□□json□jsonb□□□□□□□□□jsonb□□□□□□□json□jsonb□□□□□□□□□□□□json□jsonb□□□□□□□□

9.1 □jsonb□□□□□

□□□□□□jsonb□□□□□□□□□jsonb□□□□□□GIN□□□□□□□□□□□□□□json□□□□□□□□□□□□jsonb□□□□□

```
{
    "id": 1,
    "user_id": 1440933,
    "user_name": "1_francs",
    "create_time": "2017-08-03 16:22:05.528432+08"
}
```

□□□□□□jsonb□□□□□□□user_info□□□□tbl_user_jsonb□□user_info□□□□□GIN□□□□□□□□□□

```
CREATE INDEX idx_gin ON tbl_user_jsonb USING gin(user_info);
```

jsonb□□GIN□□□□“@>”“□”“□&”“□|”□□□□□□□□□□□□□□□□□

```
SELECT * FROM tbl_user_jsonb WHERE user_info @>
'{"user_name": "1_frans"}'
```

□□□□□□jsonb□□□□□□□□□□idx_gin□

```
SELECT * FROM tbl_user_jsonb WHERE user_info->>'user_name'=
'1_francs';
```

[illegible]jsonb[illegible]jsonb[illegible]

```
CREATE INDEX idx_gin_user_infob_user_name ON tbl_user_jsonb
USING btree
((user_info ->> 'user_name'));
```

[illegible]user_info->>'user_name'[illegible]SQL[illegible]

9.2 json[illegible]jsonb[illegible]

[illegible]jsonb[illegible]json[illegible]jsonb[illegible]3[illegible]json[illegible]jsonb[illegible]json[illegible]jsonb[illegible]jsonb[illegible]json[illegible]jsonb[illegible]json[illegible]json[illegible]json[illegible]jsonb[illegible]jsonb[illegible]

9.2.1 [illegible]json[illegible]jsonb[illegible]

[illegible]json[illegible]jsonb[illegible]

·user_ini[illegible]200[illegible]

·tbl_user_json[illegible]json[illegible]200[illegible]

·tbl_user_jsonb[illegible]jsonb[illegible]200[illegible]

[illegible]user_ini[illegible]200[illegible]

```
mydb=> CREATE TABLE user_ini(id int4 ,user_id int8, user_name
character
```

```
varying(64),create_time timestamp(6) with time zone default
clock_timestamp());
CREATE TABLE

mydb=> INSERT INTO user_ini(id,user_id,user_name)
SELECT r,round(random()*2000000), r || '_francs'
FROM generate_series(1,2000000) as r;
INSERT 0 2000000
```

□□□□user_ini□□□□□json□jsonb□□□□□user_ini_json□user_ini_jsonb□□□□□□□

```
mydb=> CREATE TABLE tbl_user_json(id serial, user_info json);
CREATE TABLE
mydb=> CREATE TABLE tbl_user_jsonb(id serial, user_info
jsonb);
CREATE TABLE
```

9.2.2 json□jsonb□□□□□□

□□user_ini□□□□□row_to_json□□□□user_ini_json□□200□json□□□□□□□□

```
mydb=> \timing
    Timing is on.

mydb=> INSERT INTO tbl_user_json(user_info) SELECT
row_to_json(user_ini)
FROM user_ini;
INSERT 0 2000000
Time: 13825.974 ms (00:13.826)
```

□□□□□□□□□tbl_user_json□□200□□□□□□13□□□□□□□□user_ini□□□□□200□jsonb□□□□□□tbl_user_jsonb□□□□□□

```
    mydb=> INSERT INTO tbl_user_jsonb(user_info)
        SELECT row_to_json(user_ini)::jsonb FROM user_ini;
INSERT 0 2000000
Time: 20756.993 ms (00:20.757)
```

□□□□□□□□□tbl_user_jsonb□□□200□jsonb□□□□20□□□□□□□□□json□□□□□jsonb□□□□□□□□□□□□□□□□□□

```
mydb=> \dt+ tbl_user_json
                       List of relations
 Schema |     Name      | Type  | Owner  |  Size  |
Description
--------+---------------+-------+--------+--------+-----------
--
 pguser | tbl_user_json | table | pguser | 281 MB |
(1 row)

mydb=> \dt+ tbl_user_jsonb
                        List of relations
 Schema |      Name      | Type  | Owner  |  Size  |
Description
--------+----------------+-------+--------+--------+----------
---
 pguser | tbl_user_jsonb | table | pguser | 333 MB |
(1 row)
```

□□□□□□□□□□□□□□jsonb□□□□□□□□□json□□□

□□tbl-user-json□□□□□□□□□□□□□□

```
mydb=> SELECT * FROM tbl_user_json LIMIT 1;
    id         |
user_info
--------------+----------------------------------------------------
------------------
    2000001 |
{"id":1,"user_id":1182883,"user_name":"1_francs","create_time"
: "2017-08-03T20:59:27.42741+08:00"}
(1 row)
```

9.2.3 json□jsonb□□□□□□

□□json□jsonb□□□□□□□□□□□json□jsonb□□□□□□□□□□□□□user_info□□□user_name□□□□□□□□□□□

```
mydb=> EXPLAIN ANALYZE SELECT * FROM tbl_user_jsonb WHERE
user_info->>'user_name'='1_francs';
QUERY PLAN
--------------------------------------------------------------
------------------
 Seq Scan on tbl_user_jsonb  (cost=0.00..72859.90 rows=10042
width=143) (actual time=0.023..524.843 rows=1 loops=1)
        Filter: ((user_info ->> 'user_name'::text) =
'1_francs'::text)
        Rows Removed by Filter: 1999999
    Planning time: 0.091 ms
    Execution time: 524.876 ms
(5 rows)
```

□□SQL□□□□□524□□□□□□□user_info□□□user_name□□□□btree□□□□□□□□

```
mydb=> CREATE INDEX idx_jsonb ON tbl_user_jsonb USING btree
((user_info->>'user_name'));
```

[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT * FROM tbl_user_jsonb WHERE
user_info->>'user_name'='1_francs';
                                   QUERY PLAN
-------------------------------------------------------------------
------------------
    Bitmap Heap Scan on tbl_user_jsonb  (cost=155.93..14113.93
rows=10000 width=143) (actual time=0.027..0.027 rows=1
loops=1)
        Recheck Cond: ((user_info ->> 'user_name'::text) =
'1_francs'::text)
        Heap Blocks: exact=1
        ->  Bitmap Index Scan on idx_jsonb  (cost=0.00..153.43
rows=10000 width=0) (actual time=0.021..0.021 rows=1 loops=1)
            Index Cond: ((user_info ->> 'user_name'::text) =
'1_francs'::text)
    Planning time: 0.091 ms
    Execution time: 0.060 ms
(7 rows)
```

[illegible]SQL[illegible]0.060ms[illegible]tbl_user_json[illegible]tbl_user_jsonb[illegible]user_info[illegible]id[illegible]

```
mydb=> CREATE INDEX idx_gin_user_info_id ON tbl_user_json
USING btree
(((user_info ->> 'id')::integer));
CREATE INDEX

mydb=> CREATE INDEX idx_gin_user_infob_id ON tbl_user_jsonb
USING btree
```

```
(((user_info ->> 'id')::integer));
CREATE INDEX
```

□□□□□□□tbl_user_json□□□□□□

```
mydb=> EXPLAIN ANALYZE SELECT id,user_info->'id',user_info->'user_name'
    FROM tbl_user_json
     WHERE (user_info->>'id')::int4>1 AND (user_info->>'id')::int4<10000;
                               QUERY PLAN
--------------------------------------------------------------------------
    Bitmap Heap Scan on tbl_user_json  (cost=166.30..14178.17 rows=10329 width=68) (actual time=1.167..26.534 rows=9998 loops=1)
        Recheck Cond: ((((user_info ->> 'id'::text))::integer > 1) AND (((user_info ->> 'id'::text))::integer < 10000))
        Heap Blocks: exact=338
        ->  Bitmap Index Scan on idx_gin_user_info_id (cost=0.00..163.72 rows=10329 width=0) (actual time=1.110..1.110 rows=19996 loops=1)
            Index Cond: ((((user_info ->> 'id'::text))::integer > 1) AND (((user_info ->> 'id'::text))::integer < 10000))
    Planning time: 0.094 ms
    Execution time: 27.092 ms
(7 rows)
```

□□□□□□□□□□□□□tbl_user_json□user_info□□id□□□1□10000□□□□□□□□□□□□□□□□27.092□□□□□□□tbl_user_jsonb□□□□□SQL□□□□□□□□□□

```
mydb=> EXPLAIN ANALYZE SELECT id,user_info->'id',user_info->'user_name'
    FROM tbl_user_jsonb
```

```
WHERE (user_info->>'id')::int4>1 AND (user_info-
>>'id')::int4<10000;
QUERY PLAN
-------------------------------------------------------------
------------------
    Bitmap Heap Scan on tbl_user_jsonb  (cost=158.93..14316.93
rows=10000 width=68) (actual time=1.140..8.116 rows=9998
loops=1)
        Recheck Cond: ((((user_info ->> 'id'::text))::integer
> 1) AND (((user_info ->> 'id'::text))::integer < 10000))
        Heap Blocks: exact=393
        ->  Bitmap Index Scan on idx_gin_user_infob_id
(cost=0.00..156.43 rows= 10000 width=0) (actual
time=1.058..1.058 rows=18992 loops=1)
            Index Cond: ((((user_info ->>
'id'::text))::integer > 1) AND (((user_info ->>
'id'::text))::integer < 10000))
    Planning time: 0.104 ms
    Execution time: 8.656 ms
(7 rows)
```

[illegible]tbl_user_jsonb[illegible]user_info[illegible]id[illegible]1[illegible]10000[illegible]8.656[illegible]jsonb[illegible]json[illegible]

[illegible]“json[illegible]jsonb[illegible]jsonb[illegible]”[illegible]key/value[illegible]

```
SELECT * FROM tbl_user_jsonb WHERE user_info @> '{"user_name":
"2_francs"}';
```

[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT * FROM tbl_user_jsonb WHERE
user_info @> '{"user_name": "2_francs"}';
                                QUERY PLAN
--------------------------------------------------------------
-------------------
    Seq Scan on tbl_user_jsonb  (cost=0.00..67733.00 rows=2000
width=143) (actual time=0.018..582.207 rows=1 loops=1)
        Filter: (user_info @> '{"user_name":
"2_francs"}'::jsonb)
        Rows Removed by Filter: 1999999
    Planning time: 0.065 ms
    Execution time: 582.232 ms
(5 rows)
```

□□□□□□□□□□□□□□582□□□□□□
tbl_user_jsonb□□user_info□□□gin□□□□□□□□

```
mydb=> CREATE INDEX idx_tbl_user_jsonb_user_Info ON
tbl_user_jsonb USING gin
    (user_Info);
CREATE INDEX
```

□□□□□□□□□□□□□□□□□□

```
mydb=> EXPLAIN ANALYZE SELECT * FROM tbl_user_jsonb WHERE
user_info @> '{"user_name": "2_francs"}';
                                  QUERY PLAN
--------------------------------------------------------------
------------------
    Bitmap Heap Scan on tbl_user_jsonb  (cost=37.50..3554.34
rows=2000 width=143) (actual time=0.079..0.080 rows=1 loops=1)
        Recheck Cond: (user_info @> '{"user_name":
"2_francs"}'::jsonb)
        Heap Blocks: exact=1
        ->  Bitmap Index Scan on idx_tbl_user_jsonb_user_info
(cost=0.00..37.00 rows=2000 width=0) (actual time=0.069..0.069
rows=1 loops=1)
            Index Cond: (user_info @> '{"user_name":
```

```
"2_francs"}'::jsonb)
    Planning time: 0.094 ms
    Execution time: 0.114 ms
(7 rows)
```

[illegible]0.114[illegible]

[illegible]json[illegible]jsonb[illegible]json[illegible]jsonb[illegible]jsonb[illegible]

9.3 [illegible]json[illegible]jsonb[illegible]

[illegible]jsonb[illegible]json[illegible]jsonb[illegible]PostgreSQL10[illegible]json[illegible]jsonb[illegible]PostgreSQL[illegible]json[illegible]jsonb[illegible]

9.3.1 PostgreSQL[illegible]

[illegible]SQL[illegible]Sphinx[illegible]PostgreSQL[illegible]PostgreSQL[illegible]

[illegible]like[illegible]

·[illegible]

·[illegible]

PostgreSQL[illegible]PostgreSQL[illegible]

1.tsvector

tsvector[illegible]tsvector[illegible]

```
mydb=> SELECT 'Hello,cat,how are u? cat is smiling!
'::tsvector;
                     tsvector
--------------------------------------------------
    'Hello,cat,how' 'are' 'cat' 'is' 'smiling!' 'u?'
(1 row)
```

[illegible]tsvector[illegible]to_tsvector[illegible]

```
mydb=> SELECT to_tsvector('english','Hello cat,');
    to_tsvector
-------------------
    'cat':2 'hello':1
(1 row)
```

2.tsquery

tsquery[illegible]“&”[illegible]“|”[illegible]“[illegible]”[illegible]tsquery[illegible]

```
mydb=> SELECT  'hello&cat'::tsquery;
        tsquery
-----------------
```

```
    'hello' & 'cat'
(1 row)
```

□□□□□□□tsquery□□□□□□□□□□□□□□□
to_tsquery□□□□□□□□□□□□□□□□

```
mydb=> SELECT to_tsquery( 'hello&cat' );
        to_tsquery
-----------------
    'hello' & 'cat'
(1 row)
```

□□□□□□□□□□□□□□□□□□□□□□□□
□“hello”□“cat”□□□□□□□□□□□

```
mydb=> SELECT to_tsvector('english','Hello cat,how are u')
@@
to_tsquery( 'hello&cat' );
    ?column?
----------
    t
(1 row)
```

□□□□□□□□□□□“hello”□“dog”□□□□□□□□□□□
□□□□□

```
mydb=> SELECT to_tsvector('english','Hello cat,how are u')
@@
    to_tsquery( 'hello&dog' );
    ?column?
----------
```

```
     f
(1 row)
```

□□□□□□□□□□tsquery□□□□□□□□□“|”“□”□□

□□ □□□□□□□□□□to_tsvector□□□□□to_tsvector□□□□□□□□□□□to_tsvector□[config regconfig□]document text□□□□to_tsvector□□□□□config□□□english□□□□□□config□□□□□□□□default_text_search_config□□□□□□

3.□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□200□□□□□□□□□□□

```
mydb=> CREATE TABLE test_search(id int4,name text);
CREATE TABLE
mydb=> INSERT INTO test_search(id,name) SELECT n,
n||'_francs'
FROM generate_series(1,2000000) n;
INSERT 0 2000000
```

□□□□SQL□□□test_search□name□□□□□□1_francs□□□□

```
mydb=> SELECT * FROM test_search WHERE name LIKE '1_francs';
    id |   name
-------+----------
     1 | 1_francs
(1 row)
```

□□□□□□□□□

```
mydb=> EXPLAIN ANALYZE SELECT * FROM test_search WHERE name
LIKE '1_francs';
                                      QUERY PLAN
--------------------------------------------------------------
--------------------
    Seq Scan on test_search  (cost=0.00..38465.04 rows=204
width=18) (actual time=0.022..261.766 rows=1 loops=1)
        Filter: (name ~~ '1_francs'::text)
        Rows Removed by Filter: 1999999
    Planning time: 0.101 ms
    Execution time: 261.796 ms
(5 rows)
```

□□□□□□□□□□□□□□□□□□□261□□□□□□□□□□
□□□□□□□□□□□□

```
mydb=> CREATE INDEX idx_gin_search ON test_search USING gin
(to_tsvector('english',name));
CREATE INDEX
```

□□□□SQL□□□test_search□name□□□□□□
1_francs□□□□

```
mydb=> SELECT * FROM test_search WHERE
to_tsvector('english',name) @@
to_tsquery('english','1_francs');
    id |   name
-------+----------
     1 | 1_francs
(1 row)
```

[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT * FROM test_search WHERE
to_tsvector('english', name) @@
to_tsquery('english','1_francs');
QUERY PLAN
------------------------------------------------------------
--------------------
    Bitmap Heap Scan on test_search  (cost=18.39..128.38
rows=50 width=36) (actual time=0.071..0.071 rows=1 loops=1)
        Recheck Cond: (to_tsvector('english'::regconfig,
name) @@ '''1'' & ''franc'''::tsquery)
        Heap Blocks: exact=1
        ->  Bitmap Index Scan on idx_gin_search
(cost=0.00..18.38 rows=50 width=0) (actual time=0.064..0.064
rows=1 loops=1)
            Index Cond: (to_tsvector('english'::regconfig,
name) @@ '''1'' & ''franc'''::tsquery)
    Planning time: 0.122 ms
    Execution time: 0.104 ms
(7 rows)
```

[illegible]0.104[illegible]3[illegible]SQL[illegible]

```
mydb=> EXPLAIN ANALYZE SELECT * FROM test_search
        WHERE to_tsvector(name) @@ to_tsquery('1_francs');
QUERY PLAN
------------------------------------------------------------
```

```
--------------------
    Seq Scan on test_search  (cost=0.00..1037730.00 rows=50
width=18) (actual time=0.036..10297.764 rows=1 loops=1)
        Filter: (to_tsvector(name) @@
to_tsquery('1_francs'::text))
        Rows Removed by Filter: 1999999
    Planning time: 0.098 ms
    Execution time: 10297.787 ms
(5 rows)
```

□□□□□□□□□□□□to_tsvector□'english'□name□□□□□□□□□□□□□□□□where□□□□to_tsvector□□□□□□□□□□□□□□□to_tsvector□name□□□□□□

9.3.2 json□jsonb□□□□□□

□PostgreSQL10□□□□□□□□□□□□json□jsonb□□□□□10□□□□□□□□□□□□□□□□□json□jsonb□□□□□□□□□□□□□10□□□□□□□□□□

1.PostgreSQL10□□□9.6□□to_tsvector□□□□□

□□□□9.6□□□to_tsvector□□□□□□□□□

```
[postgres@pghost1 ~]$ psql francs francs
psql (9.6.3)
Type "help" for help.

mydb=> \df *to_tsvector*
                                List of functions
    Schema  |       Name        | Result data type |
```

```
Argument data types |  Type
-----------+------------------+-----------------+--------
------------+------
 pg_catalog | array_to_tsvector | tsvector         | text[]
| normal
 pg_catalog | to_tsvector       | tsvector         |
regconfig, text     | normal
 pg_catalog | to_tsvector       | tsvector         | text
| normal
(3 rows)
```

□□□□□9.6□□to_tsvector□□□□□□□□□text□text[]□□□□□□□□□10□□□to_tsvector□□□□□□□□

```
[postgres@pghost1 ~]$ psql mydb pguser
psql (10.0)
Type "help" for help.
mydb=> \df *to_tsvector*
                                List of functions
    Schema  |       Name        | Result data type |
Argument data types |  Type
-----------+------------------+-----------------+--------
------------+------
 pg_catalog | array_to_tsvector | tsvector         | text[]
| normal
 pg_catalog | to_tsvector       | tsvector         | json
| normal
 pg_catalog | to_tsvector       | tsvector         | jsonb
| normal
 pg_catalog | to_tsvector       | tsvector         |
regconfig, json     | normal
 pg_catalog | to_tsvector       | tsvector         |
regconfig, jsonb    | normal
 pg_catalog | to_tsvector       | tsvector         |
regconfig, text     | normal
 pg_catalog | to_tsvector       | tsvector         | text
| normal
(7 rows)
```

[illegible]10[illegible]to_tsvector[illegible]json[illegible]jsonb[illegible]

2.[illegible]

[illegible]random_range[illegible]int4[illegible]int4[illegible]

```
CREATE OR REPLACE FUNCTION random_range(int4, int4)
RETURNS int4
LANGUAGE SQL
AS $$
    SELECT ($1 + FLOOR(($2 - $1 + 1) * random() ))::int4;
$$;
```

[illegible]random_text_simple[illegible]length int4[illegible]random_range[illegible]int4[illegible]int4[illegible]

```
CREATE OR REPLACE FUNCTION random_text_simple(length int4)
RETURNS text
LANGUAGE PLPGSQL
AS $$
DECLARE
    possible_chars text :=
'0123456789ABCDEFGHIJKLMNOPQRSTUVWXYZ';
    output text := '';
    i int4;
    pos int4;
BEGIN

    FOR i IN 1..length LOOP
        pos := random_range(1, length(possible_chars));
```

```
        output := output || substr(possible_chars, pos, 1);
    END LOOP;

    RETURN output;
END;
$$;
```

　　random_text_simple□length int4□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT random_text_simple(3);
    random_text_simple
--------------------
    LL9
(1 row)
```

　　□□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT random_text_simple(6);
    random_text_simple
--------------------
    B81BPW
(1 row)
```

　　□□□□□□□□□□□□□□□□

3.□□json□□□

　　□□user_ini□□□□□□□random_text_simple□length int4□□□□□100□□□□□□□□□□□□□□□□□

[illegible]

```
mydb=> CREATE TABLE user_ini(id int4 ,user_id int8,
user_name character varying(64),
create_time timestamp(6) with time zone default
clock_timestamp());
CREATE TABLE

mydb=> INSERT INTO user_ini(id,user_id,user_name)
SELECT r,round(random()*1000000), random_text_simple(6)
FROM generate_series(1,1000000) as r;
INSERT 0 1000000
```

[illegible]tbl_user_search_json[illegible] row_to_json[illegible]user_ini[illegible]json[illegible]

```
mydb=> CREATE TABLE tbl_user_search_json(id serial,
user_info json);
CREATE TABLE

mydb=> INSERT INTO tbl_user_search_json(user_info)
    SELECT row_to_json(user_ini) FROM user_ini;
INSERT 0 1000000
```

[illegible]

```
mydb=> SELECT * FROM tbl_user_search_json LIMIT 1;
    id    |
user_info
----------+------------------------------------------------------------
--------------------
        1 |
{"id":1,"user_id":186536,"user_name":"KTU89H","create_time":
```

```
"2017-08-05T15:59:25.359148+08:00"}
(1 row)
```

4.json□□□□□□□□

□□□□□□□□□tbl_user_search_json□user_info□□□□□KTU89H□□□□□□□□□□□

```
mydb=> SELECT * FROM tbl_user_search_json
WHERE to_tsvector('english',user_info) @@
to_tsquery('ENGLISH','KTU89H');
    id    |
user_info
----------+------------------------------------------------------------------
        1 |
{"id":1,"user_id":186536,"user_name":"KTU89H","create_time":
"2017-08-05T15:59:25.359148+08:00"}
(1 row)
```

□□SQL□□□□□□□□□□□□□json□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□8061□□□□□□□□

```
mydb=> EXPLAIN ANALYZE SELECT * FROM tbl_user_search_json
        WHERE to_tsvector('english',user_info) @@
to_tsquery('ENGLISH','KTU89H');
QUERY PLAN
----------------------------------------------------------------------------------------
    Seq Scan on tbl_user_search_json  (cost=0.00..279513.00
rows=5000 width=104) (actual time=0.046..8061.858 rows=1
loops=1)
        Filter: (to_tsvector('english'::regconfig,
user_info) @@ '''ktu89h''':: tsquery)
        Rows Removed by Filter: 999999
```

```
    Planning time: 0.091 ms
    Execution time: 8061.880 ms
(5 rows)
```

□□□□□□□

```
mydb=>  CREATE INDEX idx_gin_search_json ON
tbl_user_search_json USING
gin(to_tsvector('english',user_info));
    CREATE INDEX
```

□□□□□□□□□□□□SQL□□□□□□

```
mydb=> EXPLAIN ANALYZE SELECT * FROM tbl_user_search_json
WHERE to_tsvector('english', user_info) @@
to_tsquery('ENGLISH','KTU89H');
QUERY PLAN
------------------------------------------------------------
--------------------
    Bitmap Heap Scan on tbl_user_search_json
(cost=50.75..7876.06 rows=5000 width=104) (actual
time=0.024..0.024 rows=1 loops=1)
        Recheck Cond: (to_tsvector('english'::regconfig,
user_info) @@ '''ktu89h''':: tsquery)
        Heap Blocks: exact=1
        ->  Bitmap Index Scan on idx_gin_search_json
(cost=0.00..49.50 rows=5000 width=0) (actual
time=0.018..0.018 rows=1 loops=1)
            Index Cond: (to_tsvector('english'::regconfig,
user_info) @@ '''ktu89h''':: 
              tsquery)
    Planning time: 0.113 ms
    Execution time: 0.057 ms
(7 rows)
```

□□□□□□□□□□□□□□□□□□□□□□0.057□□□□□□□□□□

□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL10□□□□□□□□□□□□□json□jsonb□□□□□□

9.4 [illegible]

[illegible]PostgreSQL[illegible]NoSQL[illegible]jsonb[illegible]json[illegible]jsonb[illegible]json[illegible]PostgreSQL[illegible]json[illegible]jsonb[illegible]PostgreSQL10[illegible]json[illegible]jsonb[illegible]PostgreSQL[illegible]PostgreSQL[illegible]

□□□

[illegible]10[illegible] [illegible]

[illegible]CPU[illegible] I/O[illegible] PostgreSQL[illegible]SQL[illegible]

[illegible]PostgreSQL[illegible]PostgreSQL[illegible]SQL[illegible]OLTP[illegible]

10.1 [illegible]

[illegible]CPU[illegible]

[illegible]I/O[illegible]SATA SSD[illegible]PCIe SSD[illegible]SSD[illegible]SSD[illegible]SSD[illegible]SSD[illegible]SSD[illegible]SAN[illegible]NAS[illegible]SAN[illegible]NAS[illegible]
NFS[illegible]

CPU[illegible]
CPU[illegible]CPU[illegible]PostgreSQL[illegible]BIOS[illegible]CPU[illegible]CPU[illegible]

[illegible] I/O [illegible]
CPU [illegible] I/O [illegible]
[illegible]

10.2 [illegible]

[illegible]DBA[illegible]

10.2.1 [illegible]Linux[illegible]

Linux[illegible]CPU[illegible]I/O[illegible]Linux[illegible]top[illegible]free[illegible]vmstat[illegible]iostat[illegible]mpstat[illegible]sar[illegible]pidstat[illegible]top[illegible]free[illegible]sysstat[illegible]

[illegible]CentOS[illegible]sysstat[illegible]

```
[root@pghost1 ~]# yum install -y sysstat
```

1.top

top[illegible]top[illegible]

□□□□□□□□□□□□□□□□□□top□□□□□□□□□□

```
top - 12:01:03 up 93 days, 23:30,  1 user,  load average:
1.08, 1.09, 1.08
Tasks: 1042 total,   3 running, 1039 sleeping,   0 stopped,
0 zombie
Cpu(s):  5.3%us,  1.6%sy,  0.0%ni, 92.5%id,  0.2%wa,
0.0%hi,  0.3%si,  0.0%st
Mem:  330604220k total, 323235920k used,  7368300k free,
248996k buffers
Swap: 67108860k total,        0k used, 67108860k free,
295053464k cached
        PID USER      PR  NI  VIRT  RES  SHR S %CPU %MEM
TIME+  COMMAND
145138 postgres  20   0 37.7g  31g  31g S 26.2 10.1
5:38.22 postgres: pguser mydb 127.0.0.1(50382)
    58327 pgbounce  20   0 49368 8808  864 S 17.4  0.0
14725:14 /usr/bin/pgbouncer -d -q
/etc/pgbouncer/pgbouncer.ini
183682 postgres  20   0 37.6g  22g  21g R 17.3  7.0
1:49.69 postgres: pguser mydb 127.0.0.1(60026)
182679 postgres  20   0 37.7g  23g  22g S 15.3  7.4
2:11.07 postgres: pguser mydb 127.0.0.1(59112)
    58123 postgres  20   0 37.1g  36g  36g R 13.6 11.7
13623:27 postgres: startup process
164415 postgres  20   0 37.7g  26g  26g S 12.2  8.6
2:51.10 postgres: pguser mydb 127.0.0.1(42440)
    10674 root      20   0     0    0    0 S 11.7  0.0
22882:14 [shn_comp_wqa]
164421 postgres  20   0 37.7g  29g  28g S 11.6  9.2
3:55.88 postgres: pguser mydb 127.0.0.1(42452)
    57570 root      20   0     0    0    0 S  6.0  0.0
6386:38 [flush-252:0]
    73195 postgres  20   0 37.1g 3532 2280 S  6.0  0.0
6207:28 postgres: wal receiver process
    58145 postgres  20   0 37.1g  36g  36g S  3.6 11.6
4936:07 postgres: writer process
148192 postgres  20   0 37.7g  32g  31g S  3.4 10.3
5:57.01 postgres: pguser mydb 127.0.0.1(53174)
    10683 root      20   0     0    0    0 S  2.1  0.0
3233:57 [shn_handle_luna]
164413 postgres  20   0 37.7g  28g  27g S  2.1  9.0
3:31.48 postgres: pguser mydb 127.0.0.1(42436)
```

```
    10681 root      20   0     0    0    0 S  1.7  0.0
1579:57 [shn_gc_wqa]
    10675 root      20   0     0    0    0 S  1.1  0.0
886:09.06 [shn_wqa]
    73174 postgres  20   0  184m 5816 1036 S  1.0  0.0
745:02.12 postgres: stats collector process
    58144 postgres  20   0 37.2g  36g  36g S  0.3 11.6
812:08.97 postgres: checkpointer process
    ...
    ...
    ...
```

top[illegible]

[illegible]

```
top - 12:01:03 up 93 days, 23:30,  1 user,  load average:
1.08, 1.09, 1.08
Tasks: 1042 total,   3 running, 1039 sleeping,   0 stopped,
0 zombie
Cpu(s):  5.3%us,  1.6%sy,  0.0%ni, 92.5%id,  0.2%wa,
0.0%hi,  0.3%si,  0.0%st
Mem:  330604220k total, 323235920k used,  7368300k free,
248996k buffers
Swap: 67108860k total,        0k used, 67108860k free,
295053464k cached
```

[illegible]

```
top - [illegible] 12:01:03 [illegible]93[illegible]23[illegible]30[illegible],[illegible]1[illegible],[illegible]
      1[illegible]5[illegible]15[illegible]: 1.08, 1.09, 1.08
[illegible]:[illegible]1042[illegible],3[illegible], 1039[illegible],0[illegible],0[illegible]
CPU:[illegible]CPU[illegible]5.3%,[illegible]CPU[illegible]1.6%,[illegible]CPU[illegible]0.0%, [illegible]CPU[illegible]92.5%,[illegible]
[illegible]IO[illegible]CPU[illegible]0.2%,[illegible]CPU[illegible]0.0%,[illegible]CPU[illegible]0.3%,[illegible]0.0%
```

```
□□:□□330604220k,□□□323235920k,□□7368300k,buffers□□□248996k
□□□□:□□67108860k,□□□0k, □□67108860k,cache□□□295053464k
```

□□□□□□□□□□□□□□□□□□□□□□□□

·PID□□□id

·USER□□□□□□

·PR□□□□□□

·NI□□□□□□□□□□

·VIRT□□□□□□□□□□□□

·RES□□□□□□□□□□□□

·SHR□□□□□□□

·S□□□□□□D=□□□□□□□□□R=□□S=□□T=□
□/□□Z=□□□□

·%CPU□□□□□□□□□CPU□□□□□□□

·%MEM□□□□□□□□□□□□□

·TIME+□□□□□□CPU□□□□□□□1/100□

·COMMAND□□□□□□□□□

top□□□□□□□□PID□USER□PR□NI□VIRT□RES□SHR□S□%CPU□%MEM□TIME+□COMMAND□□□top□□□□□□□□□□□man top□□□□□

2.free

free□□□□□□□□□□□□□□□□□□□□□

```
[root@pghost1 ~]# free -g
             total       used       free     shared
buffers     cached
Mem:           315        306          9         36
0        278
-/+ buffers/cache:         27        287
Swap:           63          0         63
```

Mem□□□□□□□□□□□□□□□□□□□□315GB□□□□306GB□□□9GB□□□total=used+free□□□□□□36GB□buffers□□□0GB□□□□□□□□□GB□□□□□□□□□□□244MB□□cache□□□278GB□□□buffers□cache□□□□□□□□□□I/O□□□□□□□□□buffers□□□□□□□□□□□□□□□□cache□□□□□□□□□□□□□-/+buffers/cache□□□□□□□□used-buffers/cached□□□□□□□□□□□□□□□□□□□□□□□

□□□□27GB□□□□□□□free+buffers/cached□□□□□□□□□□□□□□□□□□□□□□□□□□□287GB□□□□□□□□Swap□□□□Swap□

□□free□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Linux□□□□I/O□□□□□□□□□□□□□□buffer cache□page cache□buffer cache□□□□□□□□□□□□□□□□□□□□□□□□□□□buffer cache□□□□□□□□□□□□□□page cache□□□□□□□□□□□□□□□□□cache□□□□□□□□□□□□□□□□□□□Swap□□□□□□□□□□□□□□□□□□□□□□□□□□□Swap□□□□□□□□□□□□□□□□□□□□cache□□□□□□□□

□□□□□□□□□□□□□□□□□□□□

```
[root@pghost1 ~]# sync
[root@pghost1 ~]# echo 1 > /proc/sys/vm/drop_caches
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

3.vmstat

vmstat□Linux□□□□□□□□□□□□□□□□□□□□□□□□□□□□CPU□□□□□□□vmstat□□□□□□□□

vmstat delay count[illegible]delay[illegible]count[illegible]count[illegible]delay[illegible]CTRL+C[illegible]

```
[root@pghost1 ~]# vmstat 3 5
procs -----------memory---------- ---swap-- -----io---- --
system-- -----cpu-----
    r  b swpd  free   buff  cache  si  so    bi    bo   in
cs us sy id wa st
    3  0    0 7598968 243376 292787104   0   0  1354   984
0    0  4  2 94  1  0
    1  0    0 7491112 243388 292891072   0   0 27019 45264
45526 72922  4  2 94  0  0
    0  0    0 7434112 243408 292946016   0   0 11299 32468
45675 68756  4  1 94  0  0
    2  0    0 7348156 243468 293028032   0   0 19383 67697
48760 75728  4  2 94  0  0
    2  0    0 7289800 243480 293085504   0   0 12467 28441
46016 64854  4  1 95  0  0
```

vmstat[illegible]

-r[illegible]CPU[illegible]1[illegible]

-b[illegible]

-swpd[illegible]

-free[illegible]

-buff□□□□□□buffer□□□□□□buffer cache□□□□□□□□□□□□□cache□□

-cache□□□□□□page cache□□□□□□□□page□cache□

-si/so□□□□□□□Swap□□□□Swap□□□□□□□□□□□

-bi/bo□□□□□□□□□□□□□□□□□□blocks/s□

-in□□□□□□□□□□□

-cs□□□□□□CPU□□□□□□□□

4.iostat

iostat□□□□□□□□□□□□□tty□□□□□□CD-ROM□□□/□□□□□□□□□□□□□□iostat□□□□□□□□□□□□vmstat□□□□□□□□□□□□□delay□count□□□□□□□□

```
[root@pghost1 ~]# iostat -dx /dev/dfa 5 5
Linux 2.6.32-696.el6.x86_64 (pghost1)       01/25/2018
_x86_64_        (48 CPU)
Device:         rrqm/s   wrqm/s     r/s     w/s   rsec/s
wsec/s avgrq-sz avgqu-sz   await r_await w_await  svctm
%util
dfa         0.00     0.00 3679.66 11761.85 129603.44
94094.80    14.49     0.17
```

```
    0.03    0.01    0.04   0.02  27.97
Device:         rrqm/s   wrqm/s     r/s     w/s   rsec/s
wsec/s avgrq-sz avgqu-sz   await r_await w_await  svctm
%util
dfa          0.00     0.00  807.60 9619.20 33854.40 76953.60
10.63     3.89
    0.37    0.15    0.39   0.01  14.00
Device:         rrqm/s   wrqm/s     r/s     w/s   rsec/s
wsec/s avgrq-sz avgqu-sz   await r_await w_await  svctm
%util
dfa         0.00     0.00  727.80 12904.40 29558.40 103235.20
9.74     7.39
    0.54   0.18    0.56   0.01  15.18
Device:         rrqm/s   wrqm/s     r/s     w/s   rsec/s
wsec/s avgrq-sz avgqu-sz   await r_await w_await  svctm
%util
dfa        0.00     0.00 1682.00 13761.20 52283.20 110089.60
10.51    16.64
    1.07    0.15    1.18   0.01  20.76
Device:         rrqm/s   wrqm/s     r/s     w/s   rsec/s
wsec/s avgrq-sz avgqu-sz   await r_await w_await  svctm
%util
dfa       0.00     0.00  609.80 23131.60 24737.60 185052.80
8.84    16.50
    0.69    0.24    0.70   0.01  21.62
```

[illegible]

-rrqm/s[illegible]wrqm/s[illegible]OS[illegible]

-r/s[illegible]w/s[illegible]

-rsec/s[illegible]wsec/s[illegible]

-avgrq-sz[illegible]

-avgqu-sz[illegible]

-await[illegible]

-svctm[illegible]

-%util[illegible]

5.mpstat

mpstat[illegible]CPU[illegible]

```
[root@pghost1 ~]# mpstat 5 5
Linux 2.6.32-696.el6.x86_64 (pghost1)       01/25/2018
_x86_64_        (48 CPU)
02:54:30 PM  CPU    %usr   %nice    %sys %iowait    %irq
%soft  %steal  %guest   %idle
02:54:35 PM  all    3.88    0.00    1.16    0.33    0.00
0.22    0.00    0.00   94.41
02:54:40 PM  all    4.06    0.00    0.98    0.14    0.00
0.24    0.00    0.00   94.58
02:54:45 PM  all    4.06    0.00    0.98    0.13    0.00
0.23    0.00    0.00   94.60
02:54:50 PM  all    4.15    0.00    1.37    0.37    0.00
0.24    0.00    0.00   93.88
02:54:55 PM  all    6.15    0.00    1.16    0.15    0.00
0.24    0.00    0.00   92.30
Average:     all    4.46    0.00    1.13    0.22    0.00
0.23    0.00    0.00   93.95
```

[illegible]mpstat[illegible]mpstat[illegible]CPU[illegible]CPU[illegible]

[illegible]-P n[illegible]n[illegible]core[illegible]CPU 0[illegible]5 [illegible]3[illegible]

```
[root@pghost1 ~]# mpstat -P 0 5 3
Linux 2.6.32-696.el6.x86_64 (pghost1)       01/25/2018
_x86_64_        (48 CPU)
02:57:16 PM  CPU    %usr   %nice    %sys %iowait    %irq
%soft  %steal  %guest   %idle
02:57:21 PM    0   17.54    0.00    2.02    0.81    0.00
0.40    0.00    0.00   79.23
02:57:26 PM    0   19.07    0.00    2.43    0.61    0.00
0.41    0.00    0.00   77.48
02:57:31 PM    0   18.96    0.00    3.39    1.00    0.00
0.60    0.00    0.00   76.05
Average:       0   18.52    0.00    2.62    0.81    0.00
0.47    0.00    0.00   77.58
```

[illegible]

-%usr[illegible]

-%nice[illegible]CPU[illegible]

-%sys[illegible]

-%iowait[illegible]I/O[illegible]

-%irq[illegible]CPU[illegible]

-%soft[illegible]CPU[illegible]

-%steal[illegible]%guest[illegible]

-%idle[illegible]

6.sar

sar[illegible]sar[illegible]/etc/cron.d/sysstat[illegible]10[illegible]

```
[root@pghost1 ~]# cat /etc/cron.d/sysstat
# Run system activity accounting tool every 10 minutes
*/10 * * * * root /usr/lib64/sa/sa1 1 1
```

[illegible]10[illegible]1[illegible]

sar[illegible]AM/PM[illegible]sar[illegible]LANG=C[illegible]24[illegible]

sar[illegible]

[illegible]1[illegible]CPU[illegible]

[illegible]CPU[illegible]

```
[root@pghost1 ~]# sar -q
12:00:01 AM   runq-sz  plist-sz   ldavg-1   ldavg-5  ldavg-
15
12:10:01 AM         9      1306      1.17      1.17
1.19
12:20:01 AM         4      1307      1.21      1.19
1.19
...
...
...
10:20:01 AM         9      1297      1.02      1.09
1.13
10:30:01 AM         7      1300      0.95      0.89
0.99
10:40:01 AM         8      1298      1.48      1.27
1.10
...
...
...
```

□□□□□□□□□

-runq-sz□□□□□□□□□□

-plist-sz□□□□□□

-ldavg-1□ldavg-5□ldavg-15□□□□□5□□□□15□□□□□□□□

□2□□□I/O□□

□□I/O□□□□□□□□□□

```
[root@pghost1 ~]# sar -b
12:00:01 AM       tps      rtps      wtps   bread/s
bwrtn/s
12:10:01 AM  13995.40    947.73  13047.68  35553.77
104404.64
12:20:01 AM  14389.13   1162.41  13226.72  40502.67
105837.12
12:30:01 AM  16420.53   1107.85  15312.69  43851.18
122524.99
...
...
...
01:00:01 PM  15961.31   1444.19  14517.12  52564.87
116160.35
01:10:01 PM  15327.42   1201.94  14125.48  43332.46
113027.64
...
...
...
```

□□□□□□□□□

-tps□rtps□wtps□TPS□□

-bread/s□bwrtn/s□□□□□block□□□□□□□□□□□□□block□□□□512□□□□□□□□□□□Block□□□□

□3□□□□□□□□

sar□□□□□□□□/var/log/sa/□□□□□□□sar□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□/etc/sysconfig/sysstat□□HISTORY□□□□□□□□□□□□28□□□□□/var/log/sa/□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□15□22□00□00□23□00□00□CPU□□□□□□□□□□□□□□

```
[root@pghost1 ~]# sar -q -f /var/log/sa/sa15 -s 22:00:00 -e 23:00:00
10:00:01 PM   runq-sz  plist-sz   ldavg-1   ldavg-5  ldavg-15
10:10:01 PM         5      1311      1.38      1.44      1.36
10:20:01 PM         2      1312      1.15      1.18      1.25
10:30:01 PM         5      1313      1.50      1.20      1.19
10:40:01 PM         4      1312      1.01      1.20      1.16
10:50:01 PM         6      1334      1.64      1.25      1.18
Average:            4      1316      1.34      1.25      1.23
```

sar□□□□□□□□□□□□□□Linux□□□□□□□□□□□□□□□□□□□□□□□□□□gunplot□□□□□□□□□□□□□□□□grafana□□□□□□□□□□□□□□□□□□□□□□□□□□□□

7.□□□□□□□□□

nmon□□□□□□□□□top□□□□□□□□□□□□□□□□□I/O□CPU□□□□□□□□□□□□□□□□□□□□□□□OFFICE□□□□□□□

iotop□□top□□□□□□□□□□□□I/O□□□□□□□□□
□□□□□□□I/O□□□□□□□□□I/O□□□□□□□□□□□□□□
□□□□□□□□pidstat□□□□□□□□□□

□□□□iotop□□□□□□□□□□I/O□□□□□□□□□□□□
□□□□□□□□□□□

□□□□□□□□□□□□□□

-D□□□□□□□□□□

-R□□□□□□

-S□□□□□□□□□

-T□□□□□□□□□□

-X dead□□□□□□□□□□□□□

-Z□□□□□□□□□□□□□□

□□□□□D□□□□□□□□□□□I/O□□□□□□“□□□□
□”□□□□□□□□□□□□□□□□□□□□□□□□□□

```
[root@pghost1 ~]# for x in `seq 3`; do ps -eo state,pid,cmd
| grep "^D"; echo "----"; sleep 5; done;
```

[illegible]

```
[root@pghost1 ~]# while true; do date; ps auxf | awk
'{if($8=="D") print $0;}'; sleep 1; done;
```

[illegible]pidstat-d 1[illegible]I/O[illegible]

10.2.2 Linux[illegible]I/O[illegible]

[illegible]I/O[illegible]Linux[illegible]I/O[illegible]Linux[illegible]Linux[illegible]I/O[illegible]

[illegible]

```
[root@pghost1 ~]# dmesg | grep -i scheduler
io scheduler noop registered
io scheduler anticipatory registered
io scheduler deadline registered
io scheduler cfq registered (default)
```

cfq[illegible]I/O[illegible]I/O[illegible]Linux[illegible]I/O[illegible]

noop[illegible]FIFO[illegible]I/O[illegible]SSD[illegible]

deadline[illegible]FIFO[illegible]I/O[illegible]

[illegible]sda[illegible]I/O[illegible]

```
[root@pghost1 ~]# cat /sys/block/sda/queue/scheduler
noop anticipatory deadline [cfq]
```

[illegible]sda[illegible]

[illegible]shell[illegible]I/O[illegible]

```
[root@pghost1 ~]# echo noop > /sys/block/sda/queue/scheduler
[root@pghost1 ~]# cat /sys/block/sda/queue/scheduler
[noop] anticipatory deadline cfq
```

shell[illegible]/etc/grub.conf[illegible]

10.2.3 [illegible]

[illegible]I/O[illegible] Linux[illegible]I/O[illegible]Linux[illegible] I/O[illegible]

[illegible]

```
[root@pghost1 ~]# /sbin/blockdev --getra /dev/sda
256
```

[illegible]256[illegible]16384[illegible]

```
[root@pghost1 ~]# /sbin/blockdev --setra 16384 /dev/sda
```

[illegible]

```
[root@pghost1 ~]# echo 16384
/sys/block/sda/queue/read_ahead_kb
```

[illegible]/etc/rc.local[illegible]

```
[root@pghost1 ~]# echo "/sbin/blockdev --setra 16384
/dev/dfa /dev/sda1" >> /etc/rc.local
```

```
[root@pghost1 ~]# cat /etc/rc.local
#!/bin/sh
...
...
...
# Database optimisation
/sbin/blockdev --setra 16384 /dev/dfa /dev/sda1
```

10.2.4 [illegible]

1.Swap

[illegible]Swap[illegible]Swap[illegible]OOM[illegible]

[illegible]Swap[illegible]free[illegible]

```
[postgres@pghost1 ~]$ free -g
             total       used       free     shared
buffers     cached
Mem:           378         34        344         30
0          30
-/+ buffers/cache:          3        375
Swap:           31          0         31
```

□□□□□□□□□□□□□□□Swap□□□□31GB□□□□□□0□□□□□□□□□Swap□

□□□□□vmstat□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ vmstat 5 3
procs -----------memory---------- ---swap-- -----io---- --
system-- -----cpu-----
    r  b   swpd   free   buff  cache   si   so    bi    bo
in   cs us sy id wa st
    0  0      0 361127872 390616 31882912   0   0     2   11
0    0  0  0 100  0  0
    0  0      0 361127840 390620 31882912   0   0     0   30
744  533  0  0 100  0  0
    1  0      0 361127840 390620 31882912   0   0     0   22
1161  694  0  0 100  0  0
[postgres@pghost1 ~]$
```

□vmstat□□□□□□□□□swpd□□□□0□□□□□□□□Swap□□□□□□□□□□□□Swap□

□□□□□□□□□□□□□□Swap□□□□□□□□□□□□□□□□□□Swap□□□Swap□□□□□□□□□□Swap□□□□Swap□□□□□□Swap□□□□swapoff□□□Swap□□□□swapon□□□□

```
[root@pghost1 ~]# swapoff -a
[root@pghost1 ~]#
[root@pghost1 ~]# free | grep Swap
Swap:            0          0          0
[root@pghost1 ~]#
[root@pghost1 ~]# swapon -a
[root@pghost1 ~]#
```

```
[root@pghost1 ~]# free | grep Swap
Swap:     33554428          0   33554428
```

□□Swap□□□□□□□□□□□□□□□□0□□□□Swap□□□□□□□□□□□□□□□□□Swap□□□□□□

2.□□□□

□□□□□Transparent HugePages□□□

□□□□□□□□□□□□□□□□□□□

```
[root@pghost1 ~]# cat
/sys/kernel/mm/transparent_hugepage/enabled
[always] madvise never
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
[root@pghost1 ~]# echo never >
/sys/kernel/mm/transparent_hugepage/enabled
[root@pghost1 ~]# cat
/sys/kernel/mm/transparent_hugepage/enabled
always madvise [never]
```

□□□□□□□□□□□□□□/etc/rc.local□□□□□□□□

```
if test -f /sys/kernel/mm/transparent_hugepage/enabled; then
    echo never > /sys/kernel/mm/transparent_hugepage/enabled
fi
if test -f /sys/kernel/mm/transparent_hugepage/defrag; then
    echo never > /sys/kernel/mm/transparent_hugepage/defrag
fi
```

□□□□□□□/etc/grub.conf□□kernel□□□□□transparent_hugepage=never□□□□□□□□□□□□

```
kernel /boot/vmlinuz-2.6.32-642.11.1.el6.x86_64 ro
root=UUID=c429e8ae-c35d-4bc0-a781-17bbb95a75cf nomodeset
rd_NO_LUKS  KEYBOARDTYPE=pc KEYTABLE=us LANG=en_US.UTF-8
rd_NO_MD SYSFONT=latarcyrheb-sun16 crashkernel=auto
rd_NO_LVM rd_NO_DM rhgb quiet numa=off
transparent_hugepage=never
```

3.NUMA

NUMA□□□□□□□□□□□□□CPU□local□□□□□□□□□□local□□□□□□□□□local□□□□□□□□□□□□□□□CPU□□□□□□□□□□□□□□□□□NUMA□□□□□□□□□□□□□□□□□□□□□□□□□□□NUMA□

□□NUMA□□□□□□□□□□□□□□□□□□

```
[root@pghost1 ~]# numactl --hardware
available: 2 nodes (0-1)
node 0 cpus: 0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32
```

```
34 36 38 40 42 44 46
node 0 size: 196514 MB
node 0 free: 186290 MB
node 1 cpus: 1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33
35 37 39 41 43 45 47
node 1 size: 196608 MB
node 1 free: 192336 MB
node distances:
node     0   1
    0:  10  21
    1:  21  10
```

□□□numastat□□□□□□□□

```
[root@pghost1 ~]# numastat
                           node0           node1
numa_hit                27207118        27526494
numa_miss                      0               0
numa_foreign                   0               0
interleave_hit            111148          111125
local_node              27205425        27405472
other_node                  1693          121022
```

□□numactl□□□□□□□□□□□□CPU□□□□□available□2nodes□0-1□□□□□□□□□□□□□□□□□□□□□□□

□□NUMA□□□□□□□□□□□□BIOS□□□□□□□□□□□□□□□□□□□□□□□□□□Node Interleaving□□□□□□□□□□□

Node Interleaving□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Enabled□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□Disabled□□□□□□□□□□
NUMA□□□□□□□□□□□□□□□□□□□□□□Disabled
□□□□□

□□□□□□□/etc/grub.conf□□kernel□□□□□
numa=off□□NUMA□□□□□□

```
kernel /boot/vmlinuz-2.6.32-642.11.1.el6.x86_64 ro
root=UUID=c429e8ae-c35d-4bc0-a781-17bbb95a75cf nomodeset
rd_NO_LUKS  KEYBOARDTYPE=pc KEYTABLE=us LANG=en_US.UTF-8
rd_NO_MD SYSFONT=latarcyrheb-sun16 crashkernel=auto
rd_NO_LVM rd_NO_DM rhgb quiet numa=off
```

□□□□□□□NUMA□□□□□□□□□

```
[root@pghost1 ~]# numactl --hardware
available: 1 nodes (0)
node 0 cpus: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38
39 40 41 42 43 44 45 46 47
node 0 size: 393122 MB
node 0 free: 379902 MB
node distances:
node   0
    0:  10
[root@pghost1 ~]#
[root@pghost1 ~]# numastat
                           node0
numa_hit                 3613403
numa_miss                      0
numa_foreign                   0
interleave_hit            222419
local_node               3613403
other_node                     0
```

[illegible]Linux[illegible]

10.3 □□□□□

10.3.1 □□□□□□

□postgresql.conf□□

□PostgreSQL□□□shared_buffers□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□shared_buffers□□□□□□□□□□□□□□□□shared_buffers□□□□□□□□□□□□□□□□□□□□□□□□□□□shared_buffers□□□shared_buffers□□□□□□□□□□shared_buffers□□□□□□□□□□□□□□□

□□□□postgresql.conf□□shared_buffers□□□□□□□□□□PostgreSQL 10□□□□□□□□□128MB□□□PostgreSQL□□□□□□□□□□□□□shared_buffers□□□□□□□□□□□□□□□□□□□□□□□□□□□pgbench□□□□□□□□□□□□□

work_mem[illegible]hash[illegible]hash[illegible]4MB[illegible]work_mem[illegible]hash[illegible]hash[illegible]PostgreSQL[illegible]Top-N heapsort[illegible]Quick sort[illegible]External merge[illegible]External merge[illegible]work_mem[illegible]

random_page_cost[illegible]4[illegible]seq_page_cost[illegible]1.5[illegible]

PostgreSQL[illegible]

10.3.2 [illegible]

[illegible]PostgreSQL[illegible]PostgreSQL[illegible]pg_stat[illegible]pg_statio[illegible]

□□□□□□□□□□□□□□□□□□□□autovacuum□□□□□□□□□□
□□□□□□□□□□□pg_statistics□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□pg_stats□□□□□□□□□□□
□□
□□□□□□□□□□

1.pg_stat_database

□□□□□□□□□□□□□□□□pg_stat_database□□□□□
□□□□□□□□□□□□□□□□

```
mydb=# \d pg_stat_database
                     View "pg_catalog.pg_stat_database"
    Column     |           Type           | Collation |
Nullable | Default
---------------+--------------------------+-----------+-----
-----+---------
datid          | oid                      |           |
|
datname        | name                     |           |
|
numbackends    | integer                  |           |
|
xact_commit    | bigint                   |           |
|
xact_rollback  | bigint                   |           |
|
blks_read      | bigint                   |           |
|
blks_hit       | bigint                   |           |
|
tup_returned   | bigint                   |           |
|
tup_fetched    | bigint                   |           |
|
tup_inserted   | bigint                   |           |
|
```

```
tup_updated    | bigint                   |          |
|
tup_deleted    | bigint                   |          |
|
conflicts      | bigint                   |          |
|
temp_files     | bigint                   |          |
|
temp_bytes     | bigint                   |          |
|
deadlocks      | bigint                   |          |
|
blk_read_time  | double precision         |          |
|
blk_write_time | double precision         |          |
|
stats_reset    | timestamp with time zone |          |
|
```

[illegible]

·numbackends[illegible]cpu[illegible]1.5[illegible]

·blks_read[illegible]blks_hit[illegible]

·xact_commit[illegible]xact_rollback[illegible]

·deadlocks[illegible]pg_stat_reset[illegible]

[illegible]

```
SELECT blks_hit::float/(blks_read + blks_hit) as
cache_hit_ratio FROM pg_stat_database WHERE
datname=current_database();
```

[illegible]I/O[illegible]1[illegible]shared_buffers[illegible]99%[illegible]

[illegible]

```
SELECT xact_commit::float/(xact_commit + xact_rollback) as
successful_xact_ratio FROM pg_stat_database WHERE
datname=current_database();
```

[illegible]1[illegible]

[illegible]pg_stat_database[illegible]numbackends[illegible]stats_reset[illegible]stats_reset[illegible]pg_stat_reset[illegible]pg_stat_reset[illegible]“stat”“reset”[illegible]pg_statistics[illegible]ANALYZE[illegible]pg_stat_reset[illegible]

2.pg_stat_user_tables

□□□□□□□□□□□□pg_stat_user□all□_tables
□□□□□□□□□□

```
mydb=# \d pg_stat_user_tables
                  View "pg_catalog.pg_stat_user_tables"
    Column           |           Type           | Collation |
Nullable | Default
--------------------+--------------------------+-----------
+----------+---------
relid               | oid                      |           |
|
schemaname          | name                     |           |
|
relname             | name                     |           |
|
seq_scan            | bigint                   |           |
|
seq_tup_read        | bigint                   |           |
|
idx_scan            | bigint                   |           |
|
idx_tup_fetch       | bigint                   |           |
|
n_tup_ins           | bigint                   |           |
|
n_tup_upd           | bigint                   |           |
|
n_tup_del           | bigint                   |           |
|
n_tup_hot_upd       | bigint                   |           |
|
n_live_tup          | bigint                   |           |
|
n_dead_tup          | bigint                   |           |
|
n_mod_since_analyze | bigint                   |           |
|
last_vacuum         | timestamp with time zone |           |
|
```

```
last_autovacuum     | timestamp with time zone |           |
|
last_analyze        | timestamp with time zone |           |
|
last_autoanalyze    | timestamp with time zone |           |
|
vacuum_count        | bigint                   |           |
|
autovacuum_count    | bigint                   |           |
|
analyze_count       | bigint                   |           |
|
autoanalyze_count   | bigint                   |           |
|
```

last_vacuum□last_analyze□□□□□□□□□□□□□vacuum□analyze□□□□

last_autovacuum□last_autoanalyze□□□□□□□□□□autovacuum□□□□□□autovacuum□analyze□□□□

idx_scan□idx_tup_fetch□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

seq_scan□seq_tup_read□□□□□□□□□□□□□□□□□□□□□□□□□□

n_tup_ins□n_tup_upd□n_tup_del□□□□□□□□□□□□□

n_live_tup□n_dead_tup□live tuple□dead tuple□□□□□

□□□□□□□□□□□□□□□□□□□vs□□

```
SELECT sum(idx_scan)/(sum(idx_scan) + sum(seq_scan)) as
idx_scan_ratio FROM pg_stat_all_tables WHERE
schemaname='your_schema';
SELECT relname,idx_scan::float/(idx_scan+seq_scan+1) as
idx_scan_ratio FROM pg_stat_all_tables WHERE
schemaname='your_schema' ORDER BY idx_scan_ratio ASC;
```

□□□□□□□□□□□□□1□□□

3.pg_stat_statements

□□□□□□□□□□□□pg_stat_statements□postgres□□□auto_explain□□□□

□□pg_stat_statements□□□postgresql.conf□□□□□□□□□

```
shared_preload_libraries = 'pg_stat_statements'
pg_stat_statements.track = all
```

□□□□CREATE EXTENSION□□□□□□□□□

```
mydb=# CREATE EXTENSION pg_stat_statements;
CREATE EXTENSION
```

---

pg_stat_statements□□□□□□□□

```
mydb=# \d pg_stat_statements
                    View "public.pg_stat_statements"
        Column       |       Type       | Collation |
Nullable | Default
--------------------+------------------+-----------+--------
--+---------
userid              | oid              |           |
|
dbid                | oid              |           |
|
queryid             | bigint           |           |
|
query               | text             |           |
|
calls               | bigint           |           |
|
total_time          | double precision |           |
|
min_time            | double precision |           |
|
max_time            | double precision |           |
|
mean_time           | double precision |           |
|
stddev_time         | double precision |           |
|
rows                | bigint           |           |
```

```
|
shared_blks_hit     | bigint           |           |
|
shared_blks_read    | bigint           |           |
|
shared_blks_dirtied | bigint           |           |
|
shared_blks_written | bigint           |           |
|
local_blks_hit      | bigint           |           |
|
local_blks_read     | bigint           |           |
|
local_blks_dirtied  | bigint           |           |
|
local_blks_written  | bigint           |           |
|
temp_blks_read      | bigint           |           |
|
temp_blks_written   | bigint           |           |
|
blk_read_time       | double precision |           |
|
blk_write_time      | double precision |           |
|
```

---

pg_stat_statements[illegible]CPU[illegible]3[illegible]

```
mydb=# SELECT calls,total_time/calls AS
avg_time,left(query,80) FROM pg_stat_statements ORDER BY 2
DESC LIMIT 3;
    calls  |     avg_time     |
left
-----------+------------------+------------------------
    678704 | 1084.18282038266 | SELECT id,
user_id,user_name, created_time, status FROM tbl;
    678704 | 1081.78246124378 | SELECT f.* FROM tbl_f f
INNER JOIN tbl_u u ON f.id...
```

```
        126 | 365.336761904762 | SELECT tableoid, oid, 
proname, prolang, pronargs, proargtypes, prorettype, proac
(3 rows)
```

[illegible]pg_stat_statements[illegible]pg_stat_statements_reset[illegible]pg_stat_statements[illegible]

4.[illegible]SQL[illegible]

[illegible]PostgreSQL[illegible]EXPLAIN[illegible]

```
mydb=# EXPLAIN SELECT * FROM tbl;
                       QUERY PLAN
--------------------------------------------------------
 Seq Scan on tbl  (cost=0.00..20.70 rows=1070 width=48)
(1 row)
```

[illegible]EXPLAIN[illegible]ANALYZE[illegible]

```
mydb=# EXPLAIN ANALYZE SELECT * FROM tbl;
           QUERY PLAN
------------------------------
 Seq Scan on tbl  (cost=0.00..20.70 rows=1070 width=48)
(actual time=0.021..0.023 rows=3 loops=1)
```

```
 Planning time: 0.117 ms
 Execution time: 0.058 ms
(3 rows)
```

□□□□□□□□□□□□ANALYZE□□□□□□□□□□□□□□□□INSERT□UPDATE□DELETE□CREATE TABLE AS□□EXECUTE□□ANALYZE□□□□□□□□□□□□□

```
mydb=# BEGIN;
BEGIN
mydb=# EXPLAIN ANALYZE UPDATE tbl SET ival = ival * 10 WHERE
id = 1;
                  QUERY PLAN
---------------------------------------------
 Update on tbl  (cost=0.15..8.17 rows=1 width=54) (actual
time=0.159..0.159 rows=0 loops=1)
    ->  Index Scan using tbl_pkey on tbl  (cost=0.15..8.17
rows=1 width=54) (actual time=0.046..0.047 rows=1 loops=1)
            Index Cond: (id = 1)
 Planning time: 4.237 ms
 Execution time: 0.315 ms
(5 rows)
mydb=# ROLLBACK;
ROLLBACK
```

□□□Execution time□0.315ms□□□□□□□□□□□□□□□0.315ms□□□□□□□□□□Planning time□4.237ms□□□□□□□□□□□□□□4.237ms□□□□□□□□□□□□□□□□□□□□□□□□□□□□tbl

□□□tbl_pkey□□□Index Scan□□□□□□□□□□□tbl□□□□□Update□□□□□□□□□□□□□□□□cost=0.00..xxx□□□□□□□□□□□“□□”□“□□”□□rows=xxx□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□width=1917□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ANALYZE□□□□□□□□□□□□□□□□□□□

□□ANALYZE□□□□□□□□□COSTS□BUFFERS□TIMING□FORMAT□□□□□□□□□□□□□□□□□□□□

```
mydb=# EXPLAIN (ANALYZE on, TIMING on , VERBOSE on, BUFFERS
on) SELECT * FROM tbl WHERE id = 10;
                   QUERY PLAN
-----------------------------------------------
 Index Scan using tbl_pkey on public.tbl  (cost=0.15..8.17
rows=1 width=48) (actual time=0.015..0.015 rows=0 loops=1)
    Output: id, ival, description, created_time
    Index Cond: (tbl.id = 10)
    Buffers: shared hit=1
 Planning time: 0.177 ms
 Execution time: 0.065 ms
(6 rows)
```

□□EXPLAIN□□□□ANALYZE□COSTS□BUFFERS□TIMING□VERBOSE□□

session□□log_xxx_stats□□□□□□PostgreSQL□initdb□□log_statement_stats□□□session□□□□□□

□postgresql.conf□□log_parser_stats□log_planner_stats□log_statement_stats□□□□□□□□□□□□off□□□log_parser_stats□log_planner_stats□□□□□□□□□log_statement_stats□□□□□□□□□□□□□□□□□□on□

□□parser□planner□□□□□□□□□□□□□□□□□□

```
mydb=# set client_min_messages = log;
mydb=# set log_parser_stats = on;
mydb=# set log_planner_stats = on;
```

□□EXPLAIN ANALYZE□□□□□□□□□□□□

```
mydb=# EXPLAIN ANALYZE select * from tbl limit 10;
LOG:  PARSER STATISTICS
DETAIL:  ! system usage stats:
!       0.000060 elapsed 0.000000 user 0.000000 system sec
!       [0.061990 user 0.014997 sys total]
!       0/0 [600/0] filesystem blocks in/out
!       0/0 [0/2640] page faults/reclaims, 0 [0] swaps
!       0 [0] signals rcvd, 0/0 [0/0] messages rcvd/sent
!       0/0 [55/1] voluntary/involuntary context switches
```

```
LOG:  PARSE ANALYSIS STATISTICS
DETAIL:  ! system usage stats:
!       0.000086 elapsed 0.000000 user 0.000000 system sec
!       [0.061990 user 0.014997 sys total]
!       0/0 [600/0] filesystem blocks in/out
!       0/0 [0/2640] page faults/reclaims, 0 [0] swaps
!       0 [0] signals rcvd, 0/0 [0/0] messages rcvd/sent
!       0/0 [55/1] voluntary/involuntary context switches
LOG:  REWRITER STATISTICS
DETAIL:  ! system usage stats:
!       0.000001 elapsed 0.000000 user 0.000000 system sec
!       [0.061990 user 0.014997 sys total]
!       0/0 [600/0] filesystem blocks in/out
!       0/0 [0/2640] page faults/reclaims, 0 [0] swaps
!       0 [0] signals rcvd, 0/0 [0/0] messages rcvd/sent
!       0/0 [55/1] voluntary/involuntary context switches
LOG:  PLANNER STATISTICS
DETAIL:  ! system usage stats:
!       0.000165 elapsed 0.001000 user 0.000000 system sec
!       [0.063990 user 0.014997 sys total]
!       0/0 [600/0] filesystem blocks in/out
!       0/0 [0/2640] page faults/reclaims, 0 [0] swaps
!       0 [0] signals rcvd, 0/0 [0/0] messages rcvd/sent
!       0/0 [58/1] voluntary/involuntary context switches
          QUERY PLAN
------------------------------
 Limit  (cost=0.00..0.43 rows=10 width=224) (actual
time=0.028..0.034 rows=10 loops=1)
    ->  Seq Scan on tbl  (cost=0.00..64771378.60
rows=1496764160 width=224) (actual time=0.025..0.027 rows=10
loops=1)
 Planning time: 0.198 ms
 Execution time: 0.083 ms
(4 rows)
```

□□□□□□□□□□□□□□□□□

```
mydb=# set client_min_messages = log;
mydb=# set log_parser_stats = off;
mydb=# set log_planner_stats = off;
mydb=# set log_statement_stats = on;
```

```
mydb=# EXPLAIN ANALYZE select * from tbl limit 10;
LOG:  QUERY STATISTICS
DETAIL:  ! system usage stats:
!       0.000603 elapsed 0.000000 user 0.000000 system sec
!       [0.065989 user 0.014997 sys total]
!       0/0 [600/0] filesystem blocks in/out
!       0/0 [0/2640] page faults/reclaims, 0 [0] swaps
!       0 [0] signals rcvd, 0/0 [0/0] messages rcvd/sent
!       0/0 [62/1] voluntary/involuntary context switches
                   QUERY PLAN
-----------------------------------------------
 Limit  (cost=0.00..0.43 rows=10 width=224) (actual
time=0.027..0.033 rows=10 loops=1)
    ->  Seq Scan on tbl  (cost=0.00..64771378.60
rows=1496764160 width=224) (actual time=0.026..0.026 rows=10
loops=1)
 Planning time: 0.154 ms
 Execution time: 0.082 ms
(4 rows)
```

[illegible]parser[illegible]planner[illegible]DETAIL[illegible]system usage stats[illegible]CPU[illegible]3[illegible]I/O[illegible]4[illegible]5[illegible]IPC[illegible]6[illegible]

10.3.3 [illegible]

[illegible]PostgreSQL[illegible]B-tree[illegible]Hash[illegible]GiST[illegible]SP-GiST[illegible]GIN[illegible]BRIN[illegible]Bloom[illegible]B-tree[illegible]B-tree[illegible]CREATE INDEX[illegible]B-tree[illegible]

□□Hash□□□□□□□□□□□□□□GIN□□□□□□□□□□□□□□□□□□□□□□□□□GiST□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□GiST□□B-tree□R-tree□□□□□□□PostGIS□□□□□□□□□□GiST□□□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□B-tree□GiST□GIN□BRIN□□□□□□□□□□□□□□□□□32□□□□□□

□□□

□PostgreSQL□□□CREATE INDEX□□□□□□□□CONCURRENTLY□□□□□□□□□□□□CONCURRENTLY□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□MVCC□□□□□□□□□□□□□□□□□□□□“□□□□”□□□□□□□□□□□CREATE INDEX CONCURRENTLY□□

```
mydb=# CREATE UNIQUE INDEX CONCURRENTLY ON mytbl USING
btree(id);
CREATE INDEX
```

[illegible]id[illegible]mytbl_pkey[illegible]
mytbl_id_idx[illegible]

```
mydb=# SELECT
schemaname,relname,indexrelname,pg_relation_size(indexrelid)
AS index_size,idx_scan,idx_tup_read,idx_tup_fetch FROM
pg_stat_user_indexes WHERE indexrelname IN (SELECT indexname
FROM pg_indexes WHERE schemaname = 'public' AND tablename =
'mytbl');
schemaname | relname | indexrelname | index_size | idx_scan
| idx_tup_read | idx_ tup_fetch
------------------------------------------------------------
-
 public   | mytbl | mytbl_pkey   | 223051776 | 1403532 |
1413850 |    1403532
 public   | mytbl | mytbl_id_idx | 222887936 |       0 |
0 |          0
(2 rows)
```

[illegible]

```
mydb=# BEGIN;
BEGIN
mydb=# ALTER TABLE mytbl DROP CONSTRAINT mytbl_pkey;
ALTER TABLE
mydb=# ALTER TABLE mytbl ADD CONSTRAINT mytbl_id_idx PRIMARY
KEY USING INDEX mytbl_id_idx;
ALTER TABLE
mydb=# END;
COMMIT
```

[illegible]

```
mydb=# SELECT
schemaname,relname,indexrelname,pg_relation_size(indexrelid)
AS index_size,idx_scan,idx_tup_read,idx_tup_fetch FROM
pg_stat_user_indexes WHERE indexrelname IN (SELECT indexname
FROM pg_indexes WHERE#
 schemaname | relname | indexrelname | index_size | idx_scan
| idx_tup_read | idx_tup_fetch·
-------------------------------------------------------------
--
 public     | mytbl    | mytbl_id_idx |  222887936  |    0
|    0 |        0
(1 row)
```

[illegible] pg_repack[illegible]

10.4 [illegible]

[illegible]Linux[illegible]I/O[illegible]Swap[illegible]NUMA[illegible]VACUUM[illegible]ANALYZE[illegible]VACUUM FREEZE[illegible]

[illegible]11[illegible] [illegible]pgbench

[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]pgbench[illegible]

11.1 [illegible]

[illegible]

·[illegible]CPU[illegible]

·[illegible]

·[illegible]

·[illegible]

[illegible]PostgreSQL[illegible]

11.1.1 [illegible]

[illegible]

[illegible]I/O[illegible]

[illegible]

[illegible]

[illegible]

[illegible]“[illegible]”[illegible]

11.1.2 [illegible]

[illegible]

[illegible]Throughput[illegible]RT[illegible]Latency[illegible]

[illegible]TPS[illegible]90%[illegible]1[illegible]10%[illegible]5[illegible]

11.1.3 [illegible]

[illegible]CPU[illegible]I/O[illegible]

[illegible]

[illegible]

11.2 [illegible]pgbench[illegible]

TPC[illegible]Transaction Processing Performance Council[illegible] http://www.tpc.org [illegible]TPC-A[illegible]TPC-B[illegible]TPC-C[illegible]TPC-D[illegible]TPC-E[illegible]TPC-W[illegible]TPC-C[illegible]OLTP[illegible]OLTP[illegible]TPC-E[illegible]benchmarksql[illegible]sysbench[illegible]PostgreSQL[illegible]pgbench[illegible]pgbench[illegible]TPC-B[illegible]

11.2.1 pgbench[illegible]

[illegible]pgbench[illegible]

```
transaction type: <builtin: TPC-B (sort of)>
scaling factor: 100
query mode: simple
number of clients: 1
number of threads: 1
number of transactions per client: 10
number of transactions actually processed: 10/10
latency average = 2.557 ms
tps = 391.152261 (including connections establishing)
tps = 399.368200 (excluding connections establishing)
```

transaction type□□□□□□□□□□□□□□□□

scaling factor□□pgbench□□□□□□□□□□□□
□□□□□

query mode□□□□□□□□□□□□□□simple□□□
□□extended□□□□□prepared□□□□□

number of clients□□□□□□□□□□□□□

number of threads□□□□□□□□□□□□□□□□□

number of transactions per client□□□□
□□□□□□□□□□□□□□□

number of transactions actually processed□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□

latency average□□□□□□□□□□□□□□

□□□□TPS□□□□□□□□□□□□□□□□□□TPS□□

11.2.2 □□□□□□□□□□

1.□□□□□□□

pgbench□□□□□□□4□□□pgbench_branches□pgbench_tellers□pgbench_accounts□pgbench_history□□□pgbench□□□□□□□□pgbench□□□pgbench□□□□□□□□□□□□□□

pgbench□□□□□□□□□□

```
pgbench -i [OPTION]... [DBNAME]
```

pgbench□□□□□□□□□□

```
-i, --initialize         □□□□□□□□
-F, --fillfactor=NUM     □□□□□□□□□□□□□□□□□□□□□□□□□10□100□□□□
□
                          □□□□□□□□100□□□□□100□□□□UPDATE□□□□□
□□□□
                          □□□□
-n, --no-vacuum          □□□□□□□□□VACUUM□□□
-q, --quiet              □□□□□□□□□□□□□□□□□□□100000□□□□□□□□□□
□
                          □□□□□□□□□□□5□□□□□□□□□□□□□□□□□□□□□□
                          □□□□□□□□□□□
-s, --scale=NUM          □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
                          □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
                          □1□□□□□□1□□□□□□□□□□□□□□□□□□k□□□□k□
□□
                          □□□□□□□□□□□□□□□□□□□□□
pgbench_branches         1k
pgbench_tellers          10k
```

```
pgbench_accounts          100000k
pgbench_history           0
--foreign-keys            □□□□4□□□□□□□□□□□□□
--index-tablespace=TABLESPACE
                          □□□□□□□□□□□
--tablespace=TABLESPACE   □□□□□□□□□□
--unlogged-tables         □□□4□□□□□□□UNLOGGED□
```

□□□□□□□□□□□□□□□□

```
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/pgbench -i -s 2 -F
80 -h pghost1 -p 1921 -U pguser -d mydb
creating tables...
100000 of 200000 tuples (50%) done (elapsed 0.02 s,
remaining 0.02 s)
200000 of 200000 tuples (100%) done (elapsed 0.16 s,
remaining 0.00 s)
vacuum...
set primary keys...
done.
```

2.□□pgbench□□□□□□□□

pgbench□□□□□□□□tpcb-like□simple-update□select-only□□□□□□□□□□□□□□□□□□pgbench□□□□□□□□□□□□

```
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/pgbench -b list
Available builtin scripts:
    tpcb-like
    simple-update
    select-only
```

tpcb-like□□□□□□□□□□□□SELECT□UPDATE□INSERT□□□□

```
BEGIN;
    UPDATE pgbench_accounts SET abalance = abalance + :delta
WHERE aid = :aid;
    SELECT abalance FROM pgbench_accounts WHERE aid = :aid;
    UPDATE pgbench_tellers SET tbalance = tbalance + :delta
WHERE tid = :tid;
    UPDATE pgbench_branches SET bbalance = bbalance + :delta
WHERE bid = :bid;
    INSERT INTO pgbench_history (tid, bid, aid, delta,
mtime) VALUES (:tid, :bid, :aid, :delta, CURRENT_TIMESTAMP);
END;
simple-update□□□□□□□□□□
BEGIN;
    UPDATE pgbench_accounts SET abalance = abalance + :delta
WHERE aid = :aid;
    SELECT abalance FROM pgbench_accounts WHERE aid = :aid;
    INSERT INTO pgbench_history (tid, bid, aid, delta,
mtime) VALUES (:tid, :bid, :aid, :delta, CURRENT_TIMESTAMP);
END;
select-only□□□□□□□□□□
BEGIN;
    SELECT abalance FROM pgbench_accounts WHERE aid = :aid;
END;
```

□□□□□□□□□□□□□□□□□□□□

```
-b scriptname[@weight]
--builtin = scriptname[@weight]
```

---

scriptname□□□□□□□□□□□□□□□□□□□□tpcb-like□□□□□□□□□□□□simple-update□□□

[illegible]

```
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/pgbench -b simple-
update -h pghost1 -p 1921 -U pguser mydb
starting vacuum...end.
transaction type: <builtin: simple update>
scaling factor: 100
query mode: simple
number of clients: 1
number of threads: 1
number of transactions per client: 10
number of transactions actually processed: 10/10
latency average = 1.631 ms
tps = 613.238093 (including connections establishing)
tps = 631.101768 (excluding connections establishing)
```

[illegible]3[illegible]@[illegible]@[illegible]simple-update[illegible]select-only[illegible]2[illegible]8[illegible]

```
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/pgbench -b simple-
update@2 -b select-only@8 -b tpcb@0 -h pghost1 -p 1921 -U
pguser mydb
starting vacuum...end.
transaction type: multiple scripts
scaling factor: 100
query mode: simple
number of clients: 1
number of threads: 1
number of transactions per client: 10
number of transactions actually processed: 10/10
latency average = 0.617 ms
tps = 1621.779592 (including connections establishing)
tps = 1753.156910 (excluding connections establishing)
SQL script 1: <builtin: TPC-B (sort of)>
    - weight: 0 (targets 0.0% of total)
```

```
    - 0 transactions (0.0% of total, tps = 0.000000)
    - latency average = -nan ms
    - latency stddev = -nan ms
SQL script 2: <builtin: select only>
    - weight: 8 (targets 80.0% of total)
    - 9 transactions (90.0% of total, tps = 1459.601633)
    - latency average = 0.439 ms
    - latency stddev = 0.137 ms
SQL script 3: <builtin: simple update>
    - weight: 2 (targets 20.0% of total)
    - 1 transactions (10.0% of total, tps = 162.177959)
    - latency average = 1.738 ms
    - latency stddev = 0.000 ms
```

□@□□tpcb-like□□□tpcb□□□t□simple□select□□s□□□□□□□□□s□□□□□□□□si□se□□□□□□□□□□

```
ambiguous builtin name: 2 builtin scripts found for prefix
"s"
```

-b simple-update□□□□-N□--skip-some-updates□□□□□-b select-only□□□-S□--select-only□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□TPS□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

11.2.3 [illegible]

[illegible]pgbench[illegible]pgbench[illegible]

1.[illegible]

[illegible]

```
CREATE TABLE tbl
(
    id SERIAL PRIMARY KEY,
    ival INT
);
```

2.[illegible]

[illegible]pgbench[illegible]-f[illegible]

```
[postgres@pghost2 ~]$ echo "SELECT id,ival FROM tbl ORDER BY
id DESC LIMIT 10;" > bench_script_for_select.sql
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/pgbench -f
bench_script_for_select.sql -h pghost1 -p 1921 -U pguser
mydb
transaction type: bench_script_for_select.sql
...
tps = 3448.671865 (including connections establishing)
tps = 4097.559616 (excluding connections establishing)
```

[illegible]SQL[illegible]psql[illegible]pgbench[illegible]

\sleep number[us|ms|s][illegible]

\set varname expression[illegible]pgbench[illegible]random[illegible]int[illegible]int[illegible]

[illegible]tbl[illegible]ival[illegible]

```
[postgres@pghost2 ~]$ cat bench_script_for_insert.sql
\sleep 500 ms
\set ival random(1, 100000)
INSERT INTO tbl(ival) VALUES (:ival);
```

[illegible]pgbench[illegible]

```
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/pgbench -f
bench_script_for_insert.sql -h pghost1 -p 1921 -U pguser
mydb
starting vacuum...end.
transaction type: bench_script_for_insert.sql
...
latency average = 501.291 ms
number of transactions actually processed: 14846
...
```

[illegible]

```
mydb=# SELECT COUNT(*) FROM tbl;
    count
-------
    14846
(1 row)
mydb=# SELECT id,ival FROM tbl ORDER BY id DESC LIMIT 3;
     id   | ival
---------+-------
    14846 | 95018
    14845 | 88153
    14844 | 21896
(3 rows)
```

[illegible]INSERT[illegible]14846[illegible]pgbench[illegible]latency average[illegible]500[illegible]14846[illegible]

[illegible]-f[illegible]@[illegible]3[illegible]10[illegible]

```
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/pgbench -T 60 -f
bench_script_for_insert.sql@3 -f
bench_script_for_insert.sql@10 -h pghost1 -p 1921 -U pguser
mydb
starting vacuum...end.
transaction type: multiple scripts
scaling factor: 1
query mode: simple
number of clients: 1
number of threads: 1
```

```
duration: 60 s
number of transactions actually processed: 176523
latency average = 0.340 ms
tps = 2942.048332 (including connections establishing)
tps = 2942.071422 (excluding connections establishing)
SQL script 1: bench_script_for_insert.sql
    - weight: 3 (targets 23.1% of total)
    - 40921 transactions (23.2% of total, tps = 682.016280)
    - latency average = 0.340 ms
    - latency stddev = 0.035 ms
SQL script 2: bench_script_for_insert.sql
    - weight: 10 (targets 76.9% of total)
    - 135602 transactions (76.8% of total, tps =
2260.032052)
    - latency average = 0.340 ms
    - latency stddev = 0.038 ms
```

11.2.4 [illegible]

1.[illegible]

-c[illegible]1[illegible]-c[illegible]

[illegible]4[illegible]

```
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/pgbench -c 4 -h
pghost1 -p 1921 -U pguser mydb
...
number of clients: 4
...
latency average = 4.826 ms
tps = 828.781956 (including connections establishing)
tps = 894.930906 (excluding connections establishing)
```

□□□□□□□□□□□□□□□□□

```
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/pgbench -c 4 -C -h
pghost1 -p 1921 -U pguser mydb
...
number of clients: 4
...
latency average = 19.627 ms
tps = 203.797152 (including connections establishing)
tps = 256.799857 (excluding connections establishing)
```

□□□□□□CPU□□□□□□□□□□□-j□□□□□□□□□□□□□□□□□□□□

2.□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□

·-T seconds□--time=seconds□□□□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□-T 3600□□□

·-t transactions□--transactions=transactions□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□10□□

□□□

□□□□□□□□□□□□□□□pgbench□□□□□□□□□□□□□

```
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/pgbench -c 4 -h
pghost1 -p 1921 -U pguser mydb
...
number of clients: 4
...
number of transactions actually processed: 40/40
...
```

□□□□□□4□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□“□□□□□□□□□□□”□“□□□□□□□□”□40□

3.□□□□□□□□□□□

□□-R□□□□□□□□□□□□□□□TPS□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

4.□□□□□□□□□□

□□-L□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/pgbench -T 10 -L 1 -
c 8 -j 8 -f bench_script_for_select.sql@3 -f
bench_script_for_update.sql@10 -h pghost1 -p 1921 -U pguser
mydb
...
number of transactions actually processed: 129234
number of transactions above the 1.0 ms latency limit: 2226
```

```
(1.722 %)
...
```

[illegible]1[illegible]2226[illegible]1.722%[illegible]

5.[illegible]

pgbench[illegible]

[illegible]-d[illegible]debug[illegible]

[illegible]-P[illegible]2[illegible]

```
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/pgbench -P 2 -T 7200
-c 8 -j 8 -f bench_script_for_select.sql@10 -f
bench_script_for_update.sql@3 -h pghost1 -p 1921 -U pguser
mydb
...
progress: 2.0 s, 20448.0 tps, lat 0.390 ms stddev 0.160
...
progress: 8.0 s, 25129.4 tps, lat 0.318 ms stddev 0.108
progress: 10.0 s, 25064.1 tps, lat 0.319 ms stddev 0.116
...
```

-l[illegible]--log[illegible]“pgbench_log.n”[illegible]-j[illegible]“pgbench_log.n.m”[illegible]n[illegible]pgbench[illegible]PID[illegible]m[illegible]1[illegible]

pgbench_log□□□□□□□□□□□□□□□□□□□□□□□□--log-prefix prefix_name□□□□□□

```
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/pgbench -T 10 -l --
log-prefix=custom -c 6 -j 2 -f bench_script_for_insert.sql@3
-f bench_script_for_insert.sql@10 -h pghost1 -p 1921 -U
pguser mydb
```

□□□□□□□□□□pgbench□□□□□□custom.151940□custom2.151940.2□□□□□□□□□□□□□□□□□□□□

```
client_id transaction_no time script_no time_epoch time_us [
schedule_lag ]
5 17 1294 1 1515153407 106478
4 15 322 0 1515152900 291756
3 7 357 0 1515152900 291768
...
4 16 334 0 1515152900 292090
...
```

□□client_id□□□□□□□id□transaction_no□□□□□□□time□□□□□□□□□□□□□□□□□□□□□□□□□□□script_no□□□□□□□□□□□□□□time_epoch/time_us□□□Unix□□□ISO 8601□□□

11.3 [illegible]

[illegible]PostgreSQL[illegible]pgbench[illegible]pgbench[illegible]pgbench[illegible]pgbench[illegible]

[illegible]12[illegible] [illegible]

PostgreSQL[illegible]9.0[illegible]Streaming Replication[illegible]pghost1[illegible]PostgreSQL[illegible]pghost2[illegible]PostgreSQL[illegible]pghost1[illegible]Primary Database[illegible]Master[illegible]pghost2[illegible]Standby Database[illegible]Slave[illegible]pghost1[illegible]pghost2[illegible]

[illegible]Logical Replication[illegible]PostgreSQL[illegible]PostgreSQL10[illegible]Slony-I[illegible]Londiste[illegible]pglogical[illegible]PostgreSQL10[illegible]PostgreSQL 10[illegible]

WAL[illegible]Write-Ahead Logging[illegible]WAL[illegible]

[illegible]WAL[illegible]WAL[illegible]WAL[illegible]WAL[illegible]WAL[illegible]

[illegible]

·[illegible]WAL[illegible]WAL[illegible]WAL[illegible]WAL[illegible]WAL[illegible]WAL[illegible]WAL[illegible]DML[illegible]WAL[illegible]

·[illegible]PostgreSQL[illegible]

·[illegible]DDL[illegible]DDL[illegible]

·[illegible]

·[illegible]PostgreSQL[illegible]PostgreSQL[illegible]

[illegible] Delayed Standbys[illegible] Quorum Commit[illegible]

12.1 □□□□□

□□WAL□□□□□□□□□WAL□□□

□□□□□□□PostgreSQL□□pg_basebackup□□□

□□□□□□□□□□□□□□□□□12-1□

□12-1 □□□□□□□

主机	主机名	IP 地址	操作系统	PostgreSQL 版本
主节点	pghost1	192.168.28.74	CentOS6.9	PostgreSQL10
备节点	pghost2	192.168.28.75	CentOS6.9	PostgreSQL10

□□pghost1□□□□□pghost2□□□□□□□□□□□□□□□□□□□□□□□□□

12.1.1 □□□□□□□□□□□□□□

□□PostgreSQL□□□1□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□pghost1□pghost2□□□□□PostgreSQL□

□pghost1□pghost2□□□□□□□□□□□□□□□□□□□

```
# groupadd postgres
# useradd postgres -g postgres
# passwd postgres
# mkdir -p /database/pg10/pg_root
# mkdir -p /database/pg10/pg_tbs
# chown -R postgres:postgres /database/pg10
```

/database/pg10/pg_root□□□□□□□□□□□□□□/database/pg10/pg_tbs□□□□□□□□□□□□□

□□postgres□□□□□□□□□□□/home/postgres/.bash_profile□□□□□□□□□

```
export PGPORT=1921
export PGUSER=postgres
export PGDATA=/database/pg10/pg_root
export LANG=en_US.utf8
export PGHOME=/opt/pgsql
export
LD_LIBRARY_PATH=$PGHOME/lib:/lib64:/usr/lib64:/usr/local/lib
64:/lib:/usr/lib:/usr/local/lib
export PATH=$PGHOME/bin:$PATH:.
export MANPATH=$PGHOME/share/man:$MANPATH
alias rm='rm -i'
alias ll='ls -lh'
```

□□□□□□□□□□□□□□□□□□□□

□□□□□PostgreSQL□□□□□□□□□

```
# tar jxvf postgresql-10.0.tar.bz2
# cd postgresql-10.0
# ./configure --prefix=/opt/pgsql_10.0 --with-pgport=1921
```

configure□□□□□□□□□□□zlib□readline□□□□configure□□□□□□□□□□□□□□□yum install□□□□□□□□□□□□□□□□□zlib□readline□□□□□□□□□□

```
# yum install zlib readline
```

□□□□□□□□□□□□□□□

```
# gmake world
# gmake install-world
```

gmake world□□□□□□□□□□□□□□□□□□□□□□□□□□□gmake□□□□□□□□□□□□gmake world□□gmake□□□□□□□□□□□□□gmake install-world□□□□PostgreSQL□□□□□/opt/pgsql_10.0□□□□/opt/pgsql_10.0/share□□□□□□/doc□□□□□□□□/opt/pgsql_10.0/share/extension□□□□□□□□□

[illegible]PostgreSQL[illegible]gmake world[illegible]

PostgreSQL[illegible]pgsql_10.0[illegible]

```
# ln -s /opt/pgsql_10.0 /opt/pgsql
```

[illegible]PostgreSQL[illegible]root[illegible]postgres[illegible]pghost1[illegible]postgres[illegible]initdb[illegible]

```
$ initdb -D /database/pg10/pg_root -E UTF8 --locale=C -U
postgres -W
```

[illegible]/database/pg10/pg_root[illegible]$PGDATA/postgresql.conf[illegible]

```
wal_level = replica             # minimal, replica, or
logical
archive_mode = on               # enables archiving; off,
on, or always
archive_command = '/bin/date'   # command to use to archive
a logfile segment
max_wal_senders = 10            # max number of walsender
processes
wal_keep_segments = 512         # in logfile segments, 16MB
each; 0 disables
hot_standby = on
```

□□□□postgresql.conf□□□□□□□□□□□□□□□□
□□□□□□□

·wal_level□□□□WAL□□□□□□□□□□□
minimal□replica□logical□□□□□minimal□□□
WAL□□□□□□□□□□□□□□□□□□□□□□□□WAL□□□□□
□□□□□□□□□□replica□□□WAL□□□minimal□□□
□□□□□□□WAL□□□□□□□□□□□□□□□□□□□□□WAL□
□□logical□□□WAL□□□□□□□□□□□□□□□□□10□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□WAL□replica
□□□□□WAL□□□□□minimal□□□□□□logical□□□□
□WAL□□□□□replica□□□□□□□□□□□□□replica□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□replica□
□□

·archive_mode□□□□□□□□□□□off□□□□□□
□□on□□□□□□□□□archive_command□□□□□□□
□WAL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□on□

·archive_command□□□□WAL□□□□□□□□
WAL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□/bin/date□□□□□□□□□□□□□□□□□□□□□□□13□□
□□□□□□□□□□□□□□□□□□

·max_wal_senders[illegible]WAL[illegible]pg_basebackup[illegible]WAL[illegible]max_connections[illegible]10[illegible]WAL[illegible]

·wal_keep_segments[illegible]pg_wal[illegible]WAL[illegible]WAL[illegible]WAL[illegible]16MB[illegible]--with-wal-segsize[illegible]WAL[illegible]pg_wal[illegible]wal_keep_segments[illegible]×16MB[illegible]512×16MB=8GB[illegible]pg_wal[illegible]WAL[illegible]

·hot_standby[illegible]SQL[illegible]on[illegible]

[illegible]postgresql.conf[illegible]postgresql.conf[illegible]

[illegible]pg_hba.conf[illegible]

```
# replication privilege.
host    replication     repuser      192.168.28.74/32
md5
```

```
host   replication     repuser      192.168.28.75/32
md5
```

□□□□□□□□□pg_hba.conf□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□pg_hba.conf□□□□□□□

□□pghost1□□□□□□□□□□

```
[postgres@pghost1 ~]$ pg_ctl start
```

□□□□□□postgres□□□□□□□□□□□□□
repuser□□□□□□□□□REPLICATION□□□LOGIN□
□□□□□□□

```
CREATE USER repuser
    REPLICATION
    LOGIN
    CONNECTION LIMIT 5
    ENCRYPTED PASSWORD 're12a345';
```

□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□postgres□□□□□□□□□□□

```
postgres=# SELECT pg_start_backup('francs_bk1');
    pg_start_backup
```

```
----------------
    0/4000060
(1 row)
```

pg_start_backup[illegible]pghost2[illegible]

```
$ tar czvf pg_root.tar.gz pg_root --exclude=pg_root/pg_wal
$ scp pg_root.tar.gz postgres@192.168.28.75:/database/pg10
```

pg_wal[illegible]pg_wal[illegible]pg_wal[illegible]pg_root[illegible]

[illegible]pghost2[illegible]

```
$ tar xvf pg_root.tar.gz
```

[illegible]

```
postgres=# SELECT pg_stop_backup();
NOTICE:  pg_stop_backup complete, all required WAL segments
have been archived
    pg_stop_backup
----------------
    0/2000130
(1 row)
```

[illegible]
pghost2[illegible]recovery.conf[illegible]$PGDATA[illegible]$PGDATA[illegible]

```
$ cp $PGHOME/share/recovery.conf.sample
$PGDATA/recovery.conf
```

[illegible]recovery.conf[illegible]

```
recovery_target_timeline = 'latest'
standby_mode = on
primary_conninfo = 'host=192.168.28.74  port=1921
user=repuser'
```

·recovery_target_timeline[illegible]timeline[illegible]latest[illegible]latest[illegible]

·standby_mode[illegible]on[illegible]WAL[illegible]WAL[illegible]

·primary_conninfo[illegible]IP[illegible]~/.pgpass[illegible]

[illegible]~/.pgpass[illegible]

```
[postgres@pghost2 ~]$ touch .pgpass
[postgres@pghost2 ~]$ chmod 0600 .pgpass
```

.pgpass[illegible]0600[illegible].pgpass[illegible]

```
192.168.28.74:1921:replication:repuser:re12a345
192.168.28.75:1921:replication:repuser:re12a345
```

.pgpass[illegible]IP[illegible]repuser[illegible]pghost2[illegible]

```
$ pg_ctl start
```

[illegible]WAL[illegible]WAL[illegible]WAL[illegible]

```
postgres 28575 28475  0 16:41 ?        00:00:00 postgres:
wal sender process repuser 192.168.28.75(57805) streaming
0/3025000
```

[illegible]WAL[illegible]

```
postgres 15449 15331  0 16:41 ?        00:00:00 postgres:
wal receiver process   streaming 0/301FC68
```

[illegible]

```
postgres=# CREATE TABLE test_sr(id int4);
CREATE TABLE
postgres=# INSERT INTO test_sr VALUES (1);
INSERT 0 1
```

[illegible]

```
postgres=# SELECT * FROM test_sr;
    id
----
    1
(1 row)
```

[illegible]postgresql.conf[illegible]hot_standby[illegible]on[illegible]

```
hot_standby = on                        # "off" disallows
queries during recovery
```

[illegible]off[illegible]psql[illegible]

```
[postgres@pghost2 ~]$ psql postgres postgres
psql: FATAL:  the database system is starting up
```

[illegible]

[illegible]$PGDATA/pg_log[illegible]

12.1.2 [illegible]pg_basebackup[illegible]

[illegible]pg_basebackup[illegible]

1[illegible]pg_start_backup[illegible]'francs_bk1'[illegible]

2[illegible]$PGDATA[illegible]

3[illegible]pg_stop_backup[illegible]

[illegible]PostgreSQL[illegible]pg_basebackup[illegible]pg_start_backup[illegible]pg_stop_backup[illegible]pg_basebackup[illegible]

[illegible]

pg_basebackup[illegible]REPLICATION[illegible]max_wal_senders[illegible]pg_basebackup[illegible]WAL[illegible]pg_basebackup[illegible]pghost2[illegible]pg_basebackup[illegible]pghost2[illegible]

```
$ pg_ctl stop -m fast
waiting for server to shut down.... done
server stopped
$ rm -rf /database/pg10/pg_root
$ rm -rf /database/pg10/pg_tbs
```

[illegible]pghost2[illegible]pg_basebackup[illegible]

```
$ pg_basebackup -D /database/pg10/pg_root -Fp -Xs -v -P -h 
192.168.28.74 -p 1921
  -U repuser
pg_basebackup: initiating base backup, waiting for 
checkpoint to complete
pg_basebackup: checkpoint completed
pg_basebackup: write-ahead log start point: 1/B9000028 on 
timeline 1
pg_basebackup: starting background WAL receiver
7791508/7791508 kB (100%), 2/2 tablespaces
pg_basebackup: write-ahead log end point: 1/B90039E0
pg_basebackup: waiting for background process to finish
```

```
streaming ...
pg_basebackup: base backup completed
```

[illegible]pg_basebackup[illegible]checkpoint[illegible]$PGDATA[illegible]pg_basebackup[illegible]

·-D[illegible]/database/pg10/pg_root[illegible]

·-F[illegible]pg_basebackup[illegible]p[illegible]plain[illegible]t[illegible]tar[illegible]p[illegible]plain[illegible]$PGDATA[illegible]t[illegible]tar[illegible]tar[illegible]base.tar[illegible]oid.tar[illegible]OID[illegible]OID[illegible]

·-X[illegible]WAL[illegible]f[illegible]fetch[illegible]s[illegible]stream[illegible]f[illegible]fetch[illegible]WAL[illegible]wal_keep_segments[illegible]WAL[illegible]f[illegible]fetch[illegible]WAL[illegible]s[illegible]stream[illegible]WAL[illegible]WAL[illegible]WAL[illegible]

□□f□fetch□□□□□□□□WAL□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□

·-v□□□□□verbose□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□

·-P□□□□□□□□□□□□□□□□□□□□□□
pg_basebackup□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□

-h□-p□-U□□□□□□□□□□□□□□□□□□
pg_basebackup□□□□□□□□□□□□□□□
https://www.postgresql.org/docs/10/static/app-pgbasebackup.html □

pg_basebackup□□□□□□□□□□
recovery.conf□□□□□□□□□□□□□□□□□□□□□
□$PGDATA□□□□□□□□□

```
$ cp ~/recovery.conf $PGDATA
```

□□□pghost2□□□□□□□□□

```
$ pg_ctl start
```

[illegible]WAL[illegible]WAL[illegible]

12.1.3 [illegible]

[illegible]pg_stat_replication[illegible]sync_state[illegible]

```
postgres=# SELECT
usename,application_name,client_addr,sync_state
           FROM pg_stat_replication ;
    usename | application_name | client_addr | sync_state
------------+------------------+-------------+-----------
    repuser | walreceiver      | 192.168.28.75 | async
(1 row)
```

pg_stat_replication[illegible]WAL[illegible]WAL[illegible]sync_state[illegible]sync_state[illegible]

·async[illegible]

·potential[illegible]

·sync[illegible]

·quorum[illegible]PostgreSQL10[illegible]quorum standbys[illegible]

[illegible]sync_state[illegible]async[illegible]

12.2 [illegible]

[illegible]WAL[illegible]
[illegible]WAL[illegible]
WAL[illegible]

[illegible]WAL[illegible]
[illegible]WAL[illegible]

[illegible]postgresql.conf[illegible]
recovery.conf[illegible]

12.2.1 synchronous_commit[illegible]

[illegible]
synchronous_commit[illegible]
synchronous_commit[illegible]

synchronous_commit[illegible]
postgresql.conf[illegible]WAL[illegible]

□□□□□□□□□□WAL□□□□□□□□□□□□□□□□□□□□□□□□
□□on□off□local remote_apply□
remote_write□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□

·on□□□□□□□□□□□□WAL□□□WAL BUFFER□
□□WAL□□□□□□□□□on□□□□□□□□□□□□□WAL□□
WAL□□□□□□□□□□□□□□□□□on□□□□□□□□□□□□□□
□□

·off□□□□off□□□□□□□□□□□□□□WAL BUFFER
□□WAL□□□□□□□□□□□□□□□□off□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□off□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□off□

·local□local□□□□on□□□□□□□□□□□□□□□□
WAL□□□□□□□□□□□□□□

□□□□□□□□□□

·remote_write□□□□□□□□□□□□□□□□□□□□□□□
□□□□WAL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□WAL□□□□□□□□□□
□□□□□□□□□□□□□□□□WAL□□□□□□□□□WAL□□□□□

□□□□□□WAL□□□□□□□□remote_write□□□□
WAL□□□□□□□WAL□□□□□□□□□□□□□□□□□□□□□
□□□WAL□□□□□□□□□□□□□□□□□

·on□□□□on□□□□□□□□□□□□□□□□□□□□□□□
□□WAL□□□□□□WAL□□□□□□□□□□□□□□□□□□□
on□□□□WAL□□□□□□□WAL□□□□□□□□□□□□□□□
□WAL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

·remote_apply□□□□□□□□□□□□□□□□□□□□
□□□□□□□□WAL□□□WAL□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□remote_apply□□□□WAL□□
□□□□WAL□□□□□□□□□□□□□□□□□□□□□□□□□□
WAL□□□□□□□□□□□□□□□□□□□□□□□□□□□□

12.2.2 □□□□□□□

□□recovery.conf□□□□□□□□□□□□□□□□

```
primary_conninfo = 'host=192.168.28.74  port=1921
user=repuser application_name=node2'
```

primary_conninfo□□□□
application_name□□□application_name□□□
□□□□□□□□□□postgresql.conf□

synchronous_standby_names□□□□□□□application_name□□□□□□□□□□□□node2□

□□□postgresql.conf□□□□□□□□□□□□□□□□□□□□□□□□□□

```
synchronous_commit = on□ remote_apply
synchronous_standby_names = 'node2'
```

wal_level□□□□□□□□□□□□□□□□□replica□logical□□□

·synchronous_commit□□□□□on□remote_apply□□□□□□on□□□□□□□□□□WAL□□□

·synchronous_standby_names□□□node2□□□□□□□□□□□recovery.conf□□□primary_conninfo□□□application_name□□□□□□□□

□□□□□□□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ pg_ctl reload
server signaled
```

wal_level[illegible]
synchronous_commit[illegible]
synchronous_standby_names[illegible]pg_ctl reload[illegible]wal_level[illegible]replica[illegible]

[illegible]recovery.conf[illegible]pghost2[illegible]

```
[postgres@pghost2 ~]$ pg_ctl restart -m fast
waiting for server to shut down.... done
server stopped
```

[illegible]WAL[illegible]WAL[illegible]

[illegible]

```
postgres=# SELECT
usename,application_name,client_addr,sync_state
           FROM pg_stat_replication ;
    usename | application_name | client_addr   | sync_state
------------+------------------+---------------+------------
    repuser | node2            | 192.168.28.75 | sync
(1 row)
```

□□pg_stat_replication□□□sync_state□□□□□□sync□□□sync□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

12.2.3 □□□□□□□□“□□”

□□□□□□□□□□□□□□□□□□□□□□□□□□□□WAL□□“□□”□□□pghost1□□□□□□□□□pghost2□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□

```
[postgres@pghost2 ~]$ pg_ctl stop -m fast
waiting for server to shut down.... done
server stopped
```

□□□□□□□□□□□□□□□□□□□

```
postgres=# SELECT * FROM test_sr LIMIT 1;
    id
----
    1
(1 row)
```

[illegible]

[illegible]

```
postgres=# INSERT INTO test_sr(id) VALUES (5);
--[illegible]
```

[illegible]INSERT[illegible]WAL[illegible]

[illegible]PostgreSQL[illegible]12.7[illegible]

12.3 [illegible]

[illegible]PostgreSQL[illegible]
pgbench[illegible]4[illegible]CPU[illegible]8GB[illegible]

[illegible]postgresql.conf[illegible]

```
wal_level = replica            # minimal, replica, or
logical
synchronous_commit = off       # synchronization level;
```

[illegible]postgresql.conf[illegible]

```
wal_level = replica            # minimal, replica, or
logical
synchronous_commit = off       # synchronization level;
wal_keep_segments = 512        # in logfile segments, 16MB
each; 0 disables
```

[illegible]postgresql.conf[illegible]

```
wal_level = replica         # minimal, replica, or logical
synchronous_commit = on     # synchronization level;
synchronous_standby_names = 'node2'    # standby servers
that provide sync rep
```

```
wal_keep_segments = 512     # in logfile segments, 16MB
each; 0 disables
```

[illegible]postgresql.conf[illegible]postgresql.conf[illegible]

12.3.1 [illegible]

[illegible]test_per1[illegible]1000[illegible]

```
postgres=# CREATE TABLE test_per1(
    id int4,
    name text,
    create_time timestamp(0) without time zone default
clock_timestamp());
CREATE TABLE
postgres=# INSERT INTO test_per1(id,name)
    SELECT n,n||'_per1'
    FROM generate_series(1,10000000) n;
INSERT 0 10000000
```

[illegible]

```
postgres=# ALTER TABLE test_per1 ADD PRIMARY KEY(id);
ALTER TABLE
postgres=# ANALYZE test_per1;
ANALYZE
```

□□□□□□□SQL□□□□□□select_per1.sql□□
□□□□

```
\set v_id random(1,10000000)

SELECT name FROM test_per1 WHERE id=:v_id;
```

□□v_id□1□1000□□□□□□□□□□□□□□□□□□
test_per1□□□□□□□□□□□□2□4□8□16□TPS□□□
pgbench□□□□□□□□

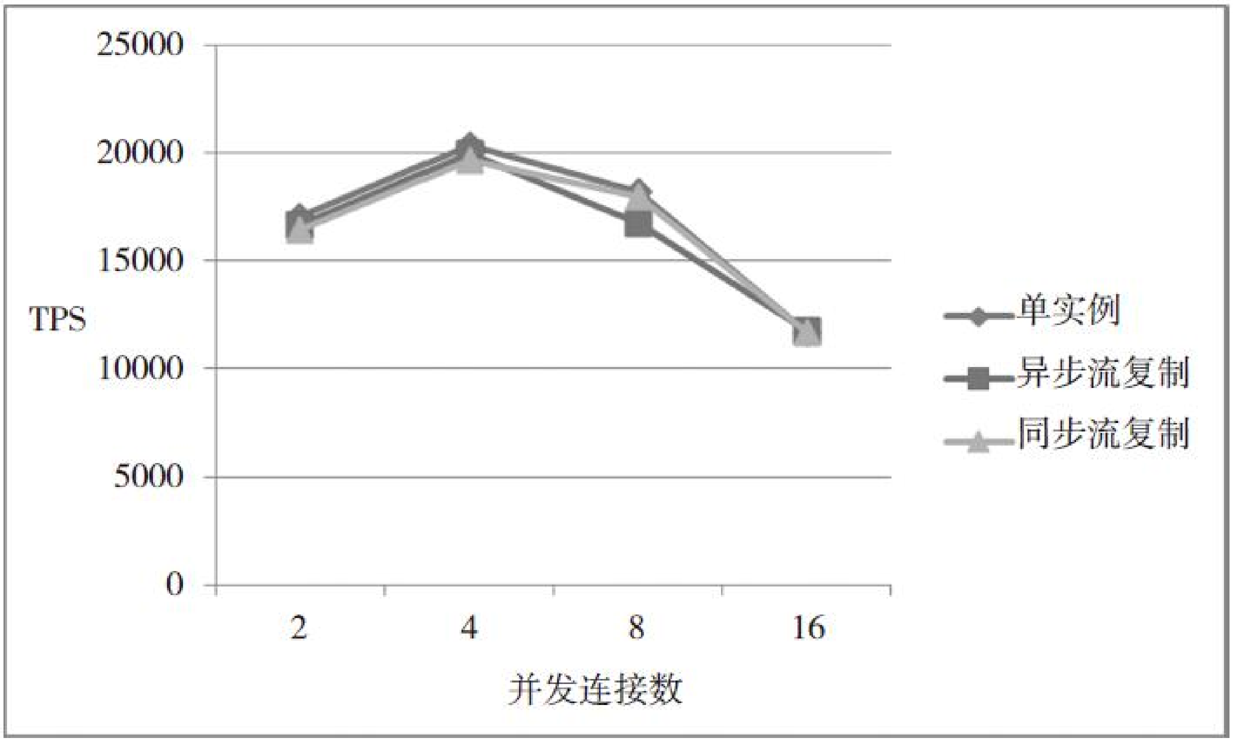
```
pgbench -c 2 -T 120 -d postgres -U postgres -n N -M prepared
-f select_per1.sql > select_2.out 2>&1 &
pgbench -c 4 -T 120 -d postgres -U postgres -n N -M prepared
-f select_per1.sql >  select_4.out 2>&1 &
pgbench -c 8 -T 120 -d postgres -U postgres -n N -M prepared
-f select_per1.sql >  select_8.out 2>&1 &
pgbench -c 16 -T 120 -d postgres -U postgres -n N -M
prepared -f select_per1.sql > select_16.out 2>&1 &
```

□□pgbench□□□□120□□-M□□repared□□□
□prepared statements□-n□□□□VACUUM□□□
□□□□pgbench□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□12-2□□□

□12-2 □□□□□□□□□□□□□□□□□□

连 接 数	单实例（TPS）	异步流复制（TPS）	同步流复制（TPS）
2	17061	16659	16427
4	20370	19966	19622
8	18193	16756	17991
16	11707	11745	11675

□□□□□□□□□□□□□□□□□□□□□12-1□□□



□12-1 □□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□4□□□□□□□□□□□□CPU□□□□□□□□□□□CPU□□□□4□□□□□□□□□□8□16□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□4□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□5%□□

□□ □□□□□□□□□CPU□□□□□□□□□□50%□□□□□□□pgbench□□□□□□□□□□□□□□□□□□□□□□□□□□□□pgbench□□□□□□□□□□□□□□□□tps□

12.3.2 □□□□□

□□test_per2□□□1000□□□□□□□□□

```
postgres=# CREATE TABLE test_per2(id int4,name text,flag
char(1));
CREATE TABLE
postgres=# INSERT INTO test_per2(id,name)
    SELECT n,n||'_per2'
    FROM generate_series(1,10000000) n;
INSERT 0 10000000
```

□□□□□□□□□□□□□□

```
postgres=# ALTER TABLE test_per2 ADD PRIMARY KEY(id);
ALTER TABLE
postgres=# ANALYZE test_per2;
ANALYZE
```

□□□□□□□□□□□□□update_per2.sql□□□□□

```
\set v_id random(1,1000000)

update test_per2 set flag='1' where id=:v_id;
```

[illegible]v_id[illegible]1[illegible]1000[illegible]test_per2[illegible]flag[illegible]2[illegible]4[illegible]8[illegible]16[illegible]TPS[illegible]pgbench[illegible]

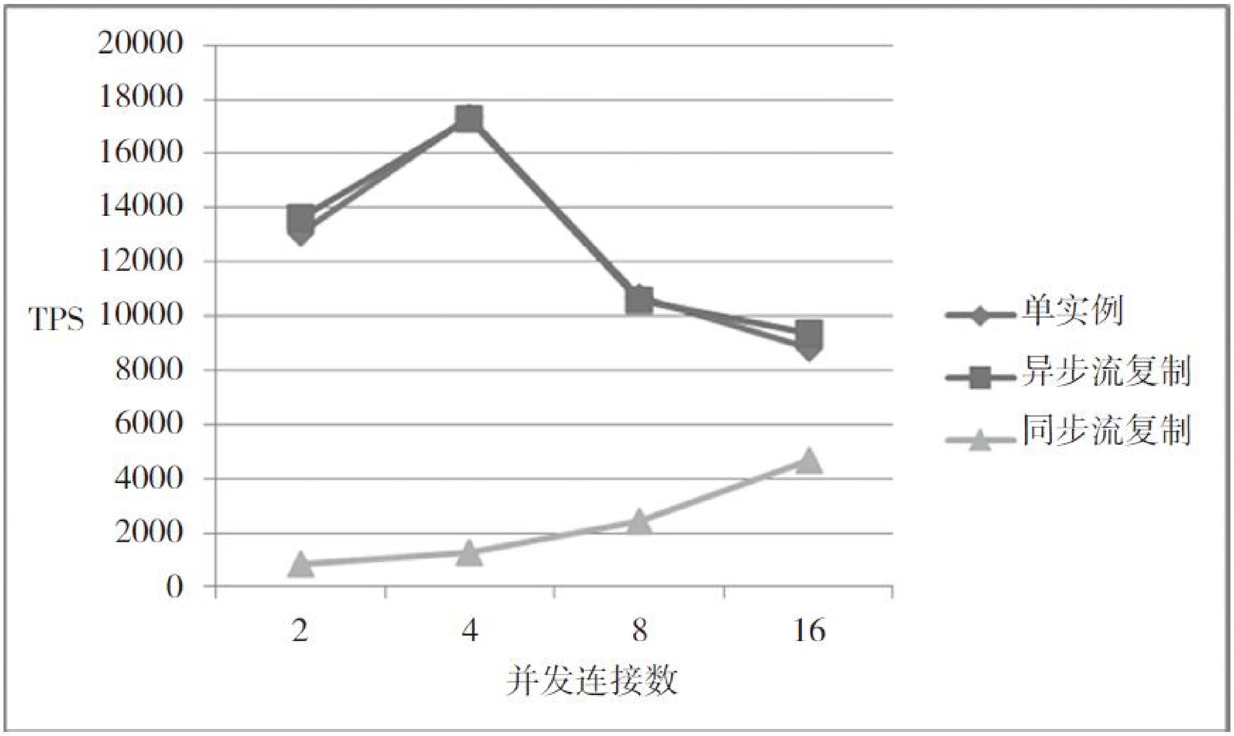
```
pgbench -c 2 -T 120 -d postgres -U postgres -n N -M prepared
-f update_per2.sql > update_2.out 2>&1 &
pgbench -c 4 -T 120 -d postgres -U postgres -n N -M prepared
-f update_per2.sql >  update_4.out 2>&1 &
pgbench -c 8 -T 120 -d postgres -U postgres -n N -M prepared
-f update_per2.sql >  update_8.out 2>&1 &
pgbench -c 16 -T 120 -d postgres -U postgres -n N -M
prepared -f update_per2.sql > update_16.out 2>&1 &
```

[illegible]pgbench[illegible]120[illegible]pgbench[illegible]12-3[illegible]

[illegible]12-3 [illegible]

连　接　数	单实例（TPS）	异步流复制（TPS）	同步流复制（TPS）
2	13062	13640	856
4	17306	17258	1270
8	10715	10598	2414
16	8793	9388	4681

[illegible]12-2[illegible]



12-2

UPDATE

WAL

TPS

12.4 [illegible]

[illegible]

12.4.1 pg_stat_replication

[illegible]WAL[illegible]pg_stat_replication[illegible]WAL[illegible]

```
postgres=# SELECT * FROM pg_stat_replication ;
-[ RECORD 1 ]----+------------------------------
pid              | 7683
usesysid         | 16384
usename          | repuser
application_name | node2
client_addr      | 192.168.28.75
client_hostname  |
client_port      | 57870
backend_start    | 2017-09-05 11:50:31.629468+08
backend_xmin     |
state            | streaming
sent_lsn         | 3/643CB568
write_lsn        | 3/643CB568
flush_lsn        | 3/643CB488
replay_lsn       | 3/643CB030
write_lag        | 00:00:00.000224
flush_lag        | 00:00:00.001562
replay_lag       | 00:00:00.006596
sync_priority    | 1
sync_state       | sync
```

□□□□□□□□□□□□

·pid□WAL□□□□□□□□

·usename□WAL□□□□□□□□□□□

·application_name□□□WAL□□□□□□□□□□□□□□□□□recovery.conf□□□□□ primary_conninfo□□application_name□□□□□

·client_addr□□□□WAL□□□□□□□IP□□□□□□□□□IP□

·backend_start□WAL□□□□□□□□□

·state□□□WAL□□□□□□□startup□□WAL□□□□□□□□catchup□□□□□□□□□□streaming□□□□□□□□□□□□□□□□□□□□WAL□□□□□□□□□□□□□□□□□backup□□□□pg_basebackup□□□□□□□stopping□□WAL□□□□□□□□

·sent_lsn□WAL□□□□□□□□WAL□□□□□

·write_lsn□□□□□□□WAL□□□□□□WAL□□□□□□□□□□□□□□□□□□WAL□□□□

·flush_lsn[illegible]WAL[illegible]WAL[illegible]WAL[illegible]

·replay_lsn[illegible]WAL[illegible]

·write_lag[illegible]WAL[illegible]WAL[illegible]WAL[illegible]WAL[illegible]

·flush_lag[illegible]WAL[illegible]WAL[illegible]WAL[illegible]WAL[illegible]WAL[illegible]

·replay_lag[illegible]WAL[illegible]WAL[illegible]WAL[illegible]WAL[illegible]WAL[illegible]

·sync_priority[illegible]quorum[illegible]

·sync_state[illegible]async[illegible]potential[illegible]sync[illegible]quorum[illegible]quorum standbys[illegible]quorum standbys[illegible]

[illegible]write_lag[illegible]flush_lag[illegible]replay_lag[illegible]PostgreSQL10[illegible]

12.4.2 [illegible]

[illegible]WAL[illegible]WAL[illegible]

[illegible]WAL[illegible]

WAL[illegible]write[illegible]flush[illegible]replay[illegible]pg_stat_replication[illegible]write_lag[illegible]flush_lag[illegible]replay_lag[illegible]WAL[illegible]SQL[illegible]

```
postgres=# SELECT
pid,usename,client_addr,state,write_lag,flush_lag,replay_lag
           FROM pg_stat_replication ;
-[ RECORD 1 ]----------------
pid         | 7683
usename     | repuser
client_addr | 192.168.28.75
state       | streaming
write_lag   | 00:00:00.000997
flush_lag   | 00:00:00.002008
replay_lag  | 00:00:00.002916
```

[illegible]WAL[illegible]WAL[illegible]WAL[illegible]write_lag[illegible]flush_lag[illegible]replay_lag[illegible]

```
replay_lag > flush_lag > write_lag
```

[illegible]flush_lag[illegible]0.2008[illegible]replay_lag[illegible]0.2916[illegible]replay_lag[illegible]flush_lag[illegible]WAL[illegible]WAL[illegible]WAL[illegible]replay_lag[illegible]flush_lag[illegible]

write_lag[illegible]flush_lag[illegible]replay_lag[illegible]PostgreSQL10[illegible]10[illegible]pg_stat_replication[illegible]SQL[illegible]

```
postgres=# SELECT EXTRACT(SECOND FROM now()-
pg_last_xact_replay_timestamp());
    date_part
-----------
    0.002227
(1 row)
```

pg_last_xact_replay_timestamp[illegible]WAL[illegible]

[illegible]WAL[illegible]
pg_last_xact_replay_timestamp[illegible]

[illegible]WAL[illegible]

[illegible]WAL[illegible]WAL[illegible]
WAL[illegible]SQL[illegible]

```
postgres=# SELECT pid,usename,client_addr,state,
        pg_wal_lsn_diff(pg_current_wal_lsn(),write_lsn)
write_delay,
        pg_wal_lsn_diff(pg_current_wal_lsn(),flush_lsn)
flush_delay,
        pg_wal_lsn_diff(pg_current_wal_lsn(),replay_lsn)
replay_dely
    FROM pg_stat_replication ;
-[ RECORD 1 ]------------
pid         | 7683
usename     | repuser
client_addr | 192.168.28.75
state       | streaming
write_delay | 560
flush_delay | 896
replay_dely | 1272
```

pg_current_wal_lsn[illegible]WAL[illegible]pg_wal_lsn_diff[illegible]WAL[illegible]WAL[illegible]write[illegible]560[illegible]flush[illegible]896[illegible]replay[illegible]1272[illegible]

[illegible]

[illegible]

12.4.3 pg_stat_wal_receiver

pg_stat_replication[illegible]WAL[illegible]WAL[illegible]

```
postgres=# SELECT * FROM pg_stat_wal_receiver ;
-[ RECORD 1 ]---------+-------------------------------------
--------------------
pid                   | 22573
status                | streaming
receive_start_lsn     | 3/2D000000
receive_start_tli     | 1
received_lsn          | 3/852DC428
received_tli          | 1
last_msg_send_time    | 2017-09-06 15:35:28.178167+08
last_msg_receipt_time | 2017-09-06 15:35:28.177706+08
latest_end_lsn        | 3/852DC508
latest_end_time       | 2017-09-06 15:35:28.178167+08
slot_name             |
conninfo               | user=repuser
passfile=/home/postgres/.pgpass dbname=replication
host=192.168.28.74 port=1921 application_name=node2
fallback_application_ name=walreceiver sslmode=disable
sslcompression=1 target_session_attrs=any
```

[illegible]

·pid WAL

·status WAL

·receive_start_lsn WAL
WAL

·received_lsn WAL WAL

·last_msg_send_time

·last_msg_receipt_time

·conninfo WAL
$PGDATA recovery.conf
primary_conninfo

12.4.4

PostgreSQL WAL

```
postgres=# SELECT pg_is_in_recovery();
    pg_is_in_recovery
-------------------
    t
(1 row)
```

□□□□□□□□□□□□□□□□□t□□□□□□□□f□□□□□□□□□□□□□□WAL□□□□□□□WAL□□□WAL□

□□□□□□□□□WAL□□□□□□□□□□

```
postgres=# SELECT pg_last_wal_receive_lsn();
    pg_last_wal_receive_lsn
-------------------------
    3/91B5BCE8
(1 row)
```

□□□□□□□□WAL□□□□□□□□□□□

```
postgres=# SELECT pg_last_wal_replay_lsn();
    pg_last_wal_replay_lsn
------------------------
    3/91EFED10
(1 row)
```

□□□□□□□□□□□□□□□□□□□

```
postgres=# SELECT pg_last_xact_replay_timestamp();
    pg_last_xact_replay_timestamp
-------------------------------
```

```
    2017-10-07 09:04:59.67741+08
(1 row)
```

□□□□WAL□□□□□□□□□□□□

```
postgres=# SELECT pg_current_wal_lsn();
    pg_current_wal_lsn
--------------------
    3/940001B0
(1 row)
```

□□□□WAL□□□□□□□□□□□□□□

```
postgres=# SELECT
pg_wal_lsn_diff('3/940001B0','3/940001A0');
    pg_wal_lsn_diff
-------------------
                 16
(1 row)
```

12.5 [illegible]

[illegible]pg_ctl promot[illegible]

[illegible]PostgreSQL[illegible]PostgreSQL[illegible]14[illegible]PostgreSQL[illegible]

12.5.1 [illegible]

[illegible]

[illegible]WAL[illegible]WAL[illegible]

[illegible]WAL[illegible]WAL[illegible]wal sender..streaming[illegible]

```
[postgres@pghost1 ~]$ ps -ef | grep "wal" | grep -v "grep"
postgres 16666 16661  0 Sep06 ?        00:00:09 postgres: wal
writer process
postgres 16672 16661  0 Sep06 ?        00:00:13 postgres: wal
sender process repuser 192.168.28.75(57872) streaming
3/9C34BCB8
```

[illegible]wal receiver..streaming[illegible]

```
[postgres@pghost2 ~]$ ps -ef | grep "wal" | grep -v "grep"
postgres 27291 22567  0 Sep06 ?        00:00:32 postgres: wal
receiver process   streaming 3/9C355788
```

[illegible]WAL[illegible]WAL[illegible]

[illegible]WAL[illegible]WAL[illegible]pg_stat_replication[illegible]

```
postgres=# SELECT
pid,usename,application_name,client_addr,state,sync_state
           FROM pg_stat_replication ;
    pid   | usename | application_name | client_addr   |
state     | sync_state
----------+---------+------------------+---------------+-----
------+------------
    16672 | repuser | node2            | 192.168.28.75 |
streaming | sync
(1 row)
```

[illegible]pg_stat_wal_receiver[illegible]

```
postgres=# SELECT
pid,status,last_msg_send_time,last_msg_receipt_time,conninfo
               FROM pg_stat_wal_receiver ;
-[ RECORD1 ]---------+--------------------------------------
-------------------
pid                   | 17551
status                | streaming
last_msg_send_time    | 2017-10-07 09:22:07.479282+08
last_msg_receipt_time | 2017-10-07 09:22:07.480277+08
conninfo         | user=repuser
passfile=/home/postgres/.pgpass dbname= replication
host=192.168.28.74 port=1921 application_name=slave1
fallback_ application_name=walreceiver sslmode=disable
sslcompression=1 target_session_attrs=any
```

[illegible]

[illegible]

```
postgres=# SELECT pg_is_in_recovery();
    pg_is_in_recovery
-------------------
    t
(1 row)
```

[illegible]t[illegible]f[illegible]

[illegible]

□□pg_controldata□□□□□□□□□□□□□□□□WAL□□□□□checkpoint□□□□□□□□□□Database cluster state□□□□□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ pg_controldata | grep cluster
Database cluster state:               in production
```

□□□□□□□□in production□□□□□□□□in archive recovery□□□□□□□□□□□

```
[postgres@pghost2 ~]$ pg_controldata | grep cluster
Database cluster state:              in archive recovery
```

□□□□□□recovery.conf□□□□□□

□□□□□□□□□□□□□□□□□$PGDATA□□□□□□recovery.conf□□□□□□□□□□□□□□□□□□□□□$PGDATA□□□□□□□□□□□□□□□□recovery.done□□□□□□□

12.5.2 □□□□□□□□□□□

PostgreSQL9.0□□□pghost1□□□□□□□□□pghost2□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

1[illegible]recovery.conf[illegible]trigger_file[illegible]

2[illegible]-m fast[illegible]

3[illegible]recovery.conf[illegible]recovery.done[illegible]

4[illegible]$PGDATA[illegible]recovery.conf[illegible]recovery.conf[illegible]$PGHOME/recovery.conf.sample[illegible]recovery.done[illegible]recovery.done[illegible]recovery.conf[illegible]primary_conninfo[illegible]IP[illegible]IP[illegible]

5[illegible]

[illegible]recovery.conf[illegible]

```
recovery_target_timeline = 'latest'
standby_mode = on
primary_conninfo = 'host=192.168.28.74  port=1921
user=repuser'
trigger_file =
'/database/pg10/pg_root/.postgresql.trigger.1921'
```

trigger_file[illegible]
[illegible]

[illegible]

```
[postgres@pghost1 ~]$ pg_ctl stop -m fast
waiting for server to shut down.... done
server stopped
```

[illegible]

```
[postgres@pghost2 pg_root]$ ll recovery.conf
-rw-r--r-- 1 postgres postgres 5.8K Sep  8 20:47
recovery.conf
[postgres@pghost2 pg_root]$ touch
/database/pg10/pg_root/.postgresql.trigger.1921
```

[illegible]recovery.conf[illegible]
trigger_file[illegible]recovery[illegible]
[illegible].conf[illegible].done[illegible]

```
    [postgres@pghost2 pg_root]$ ll recovery.done
-rw-r--r-- 1 postgres postgres 5.8K Sep  8 20:47
recovery.done
```

[illegible]

```
2017-09-08 21:00:21.622 CST,,,4357,,59b29465.1105,1,,2017-09-
08 21:00:21 CST,,0,FATAL,XX000,"could not connect to the
```

```
primary server: could not connect to server: Connection
refused
    Is the server running on host ""192.168.28.74"" and
accepting
    TCP/IP connections on port 1921?",,,,,,,,,""
2017-09-08 21:00:26.622 CST,,,4235,,59b2916f.108b,9,,2017-09-
08 20:47:43 CST,1/0,0,LOG,00000,"trigger file found:
/database/pg10/pg_root/.postgresql.trigger.1921",,,,,,,,,""
2017-09-08 21:00:26.622 CST,,,4235,,59b2916f.108b,10,,2017-
09-08 20:47:43 CST,1/0,0,LOG,00000,"redo done at
3/A0000028",,,,,,,,,""
2017-09-08 21:00:26.622 CST,,,4235,,59b2916f.108b,11,,2017-
09-08 20:47:43 CST,1/0,0,LOG,00000,"last completed
transaction was at log time 2017-09-08
20:59:30.746045+08",,,,,,,,,""
2017-09-08 21:00:26.640 CST,,,4235,,59b2916f.108b,12,,2017-
09-08 20:47:43 CST,1/0,0,LOG,00000,"selected new timeline ID:
2",,,,,,,,,""
2017-09-08 21:00:26.792 CST,,,4235,,59b2916f.108b,13,,2017-
09-08 20:47:43 CST,1/0,0,LOG,00000,"archive recovery
complete",,,,,,,,,""
2017-09-08 21:00:26.808 CST,,,4233,,59b2916f.1089,3,,2017-09-
08 20:47:43 CST,,0,LOG,00000,"database system is ready to
accept connections",,,,,,,,,""
```

[illegible]

[illegible]pg_controldata[illegible]

```
[postgres@pghost2 pg_root]$ pg_controldata | grep cluster
Database cluster state:               in production
```

[illegible]pghost2[illegible]test_alived[illegible]

```
postgres=# CREATE TABLE test_alived2(id int4);
CREATE TABLE
postgres=# INSERT INTO test_alived2 VALUES(1);
INSERT 0 1
```

□□□□□□□4□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□recovery.conf□□□□□□

```
recovery_target_timeline = 'latest'
standby_mode = on
primary_conninfo = 'host=192.168.28.75  port=1921
user=repuser'
trigger_file =
'/database/pg10/pg_root/.postgresql.trigger.1921'
```

□□□□□pghost2□□recovery.done□□□□□□□□□primary_conninfo□□□host□□□□□□□□□IP□

□□□pghost1□□postgres□□□□□□□□□~/.pgpass□□□□□□□□

```
[postgres@pghost1 ~]$ touch ~/.pgpass
[postgres@pghost1 ~]$ chmod 600 ~/.pgpass
```

□□~/.pgpass□□□□□□□□□□

```
192.168.28.74:1921:replication:repuser:re12a345
192.168.28.75:1921:replication:repuser:re12a345
```

□□□□pghost1□□□□□□□□□□□

```
[postgres@pghost1 ~]$ pg_ctl start
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□test_alived2□□□□□□□□□□

```
postgres=# SELECT * from test_alived2 ;
    id
------
    1
(1 row)
```

□□□pghost1□□□□□WAL□□□□□pghost2□□□WAL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□ □□□□□□2□□□□□□□□□□□□□□□□□□□□□□□checkpoint□□□□□□□WAL□□□□□□□□□WAL□□□□□□□□□□checkpoint□WAL□□□□□□□□□WAL□□□□□□□□□□□□□□□□□□□WAL□□□□□□WAL□□□WAL□□

12.5.3 □□□□□pg_ctl promote□□

□□□□□□□□□□□□□□□□□□□□PostgreSQL9.1□□□□□□pg_ctlpromote□□□□□□□□□□□□□□□□□□□promote□□□□□□□

```
pg_ctl promote [-D datadir]
```

-D□□□□□□□□□□□□□□□□□□□$PGDATA□□□□□promote□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

pg_ctl promote□□□□□□□□□□□□□□□□□□□□□□1□□□□□□recovery.conf□□□□□□trigger_file□□□□□□□3□□□□pg_ctl promote□□□□□□□□□□□□

1□□□□□□□□□□-m fast□□□□□

2□□□□□□□pg_ctl promote□□□□□□□□□recovery.conf□□recovery.done□□□□□□□□□□□□

3□□□□□□□□□□□□□□□□□□□□□□$PGDATA□□□□□recovery.conf□□□□□□□□□□□□recovery.conf□□□□□□□$PGHOME/recovery.conf.sample□□□□□□□□□

□□□□□□□□recovery.done□□□□□recovery.done□□□□□□recovery.conf□□□□□□□□□□□□□□primary_conninfo□□□□IP□□□□IP□

4□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□pg_ctl promote□□□□□□□□□□□□□□□□□□□□□□□□□□□pg_rewind□□□□□□□□pg_ctl promote□□□□□□□□□□

12.5.4 pg_rewind

pg_rewind□□□2□□□□□□□□□□□□□□□□□□□2□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□pghost1□□□□□□□□□pghost2□□□□□□□□□

□□recovery.conf□□□□□□□

```
recovery_target_timeline = 'latest'
standby_mode = on
primary_conninfo = 'host=192.168.28.74 port=1921
user=repuser'
```

[illegible]

```
[postgres@pghost2 pg_root]$ pg_ctl promote
waiting for server to promote.... done
server promoted
```

[illegible]pghost2[illegible]

```
[postgres@pghost2 pg_root]$ pg_controldata | grep cluster
Database cluster state:               in production
```

[illegible]pg_controldata[illegible]pghost2[illegible]pghost2[illegible]pghost1[illegible]pghost1[illegible]pghost1[illegible]

```
[postgres@pghost1 ~]$ pg_controldata | grep cluster
Database cluster state:               in production
```

pghost1[illegible]pg_controldata[illegible]pghost1[illegible]pghost2[illegible]pghost1[illegible]pghost1[illegible]

```
[postgres@pghost1 ~]$ pg_ctl stop -m fast
waiting for server to shut down.... done
server stopped
```

□$PGDATA□□□□recovery.done□□□□recovery.conf□□□□□□

```
[postgres@pghost1 pg_root]$ mv recovery.done recovery.conf
```

pghost1□recovery.conf□□□□□□□

```
recovery_target_timeline = 'latest'
standby_mode = on
primary_conninfo = 'host=192.168.28.75  port=1921
user=repuser'
```

□□pghost1□□□□□□□□□□□

```
[postgres@pghost1 pg_root]$ pg_ctl start
```

□□□□□□□□□□□□□□pghost1□□□□WAL□□□□□□pghost2□□□□□WAL□□□□□□□□□pghost1□□□□□□□□□□□□□□□□

```
2017-09-09 13:22:45.540 CST,,,11948,,59b37aa5.2eac,2,,2017-
09-09 13:22:45 CST,,0,FATAL,XX000,"could not start WAL
streaming: ERROR:  requested starting point 3/A3000000 on
timeline 3 is not in this server's history
```

```
DETAIL:  This server's history forked from timeline 3 at
3/A213F930.",,,,,,,,,""
2017-09-09 13:22:45.540 CST,,,11944,,59b37aa5.2ea8,5,,2017-
09-09 13:22:45 CST,1/0,0,LOG,00000,"new timeline 4 forked off
current database system timeline 3 before current recovery
point 3/A3000098",,,,,,,,,""
2017-09-09 13:22:45.543 CST,,,11949,,59b37aa5.2ead,1,,2017-
09-09 13:22:45 CST,,0,FATAL,XX000,"could not start WAL
streaming: ERROR:  requested starting point 3/A3000000 on
timeline 3 is not in this server's history
DETAIL:  This server's history forked from timeline 3 at
3/A213F930.",,,,,,,,,""
2017-09-09 13:22:45.543 CST,,,11944,,59b37aa5.2ea8,6,,2017-
09-09 13:22:45 CST,1/0,0,LOG,00000,"new timeline 4 forked off
current database system timeline 3 before current recovery
point 3/A3000098",,,,,,,,,"
```

[illegible]pghost1[illegible]

[illegible]PostgreSQL[illegible]pg_rewind[illegible]pg_basebackup[illegible]pg_rewind[illegible]pghost1[illegible]pghost2[illegible]

[illegible]pg_rewind[illegible]

·postgresql.conf[illegible]wal_log_hints[illegible]on[illegible]

·[illegible]initdb[illegible]--data-checksums[illegible]I/O[illegible]initdb[illegible]

[illegible]initdb[illegible]--data-checksums[illegible]postgresql.conf[illegible]wal_log_hints[illegible]

```
wal_log_hints = on
```

[illegible]

[illegible]pghost2[illegible]

```
[postgres@pghost2 pg_root]$ pg_ctl promote
waiting for server to promote.... done
server promoted
```

[illegible]pghost1[illegible]pghost1[illegible]

```
[postgres@pghost1 pg_root]$ pg_ctl stop -m fast
waiting for server to shut down.... done
server stopped
```

□□□□pg_rewind□□□□□□pghost2□□□□□pghost1□□□□□□

```
[postgres@pghost1 pg_root]$ pg_rewind --target-pgdata $PGDATA
--source-server= 'host= 192.168.28.75 port=1921 user=postgres
dbname=postgres' -P
connected to server
servers diverged at WAL location 3/A7006508 on timeline 5
rewinding from last common checkpoint at 3/A7000028 on
timeline 5
reading source file list
reading target file list
reading WAL in target
need to copy 7663 MB (total source directory size is 9237 MB)
7847309/7847309 kB (100%) copied
creating backup label and updating control file
syncing target data directory
Done!
```

□□□□□□□□□□□□□□□□postgres□□□postgres□□□□□□□~/.pgpass□□□

□recovery.done□□□□recovery.conf□□□□□□

```
[postgres@pghost1 pg_root]$ mv recovery.done recovery.conf
```

□□□□recovery.conf□primary_conninfo□□□host□□□□□□□□□□□□□□□□□□□□□□

```
[postgres@pghost1 pg_root]$ pg_ctl start
```

[illegible]pghost1[illegible]
[illegible]

12.6 [illegible]

[illegible]1[illegible]

12.6.1 [illegible]

PostgreSQL[illegible]pg_dump[illegible]

12.6.2 [illegible]

[illegible]pghost1[illegible]pghost2[illegible]

recovery_min_apply_delay□□□□□□□□ recovery.conf□□□□□□□□□□

```
recovery_min_apply_delay (integer)
```

□□□□□□□□□□□□□□□□□□□□□□□□

·ms□□□□□□□□□

·s□□□

·min□□□□

·h□□□□

·d□□□

□□□□□□□□□□□□□□□□□□□□□□□WAL□□□□□□□□□□□□□WAL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□WAL□□□□□□□□□□□□□□□□□□WAL□□□□□□□□□□□□WAL□□□WAL□□□□□□□□□□□□□□

□□□pghost2□□□□□□□□□□1□□□□□□□□

```
recovery_min_apply_delay = 1min
```

[illegible]1[illegible]

```
[postgres@pghost2 pg_root]$ pg_ctl restart
```

[illegible]test_delay[illegible]

```
post gres=# CREATE TABLE test_delay(id int4,create_time
timestamp(0) without time zone);
CREATE TABLE
```

[illegible]

```
postgres=# INSERT INTO test_delay(id,create_time) VALUES
(1,now());
INSERT 0 1
```

[illegible]test_delay[illegible]SQL[illegible]

```
postgres=# SELECT now(),create_time FROM test_delay;
-[ RECORD 1 ]-----------------------------
now         | 2017-09-10 16:18:50.414074+08
create_time | 2017-09-10 16:17:50
```

[illegible]

[illegible]

[illegible]

```
postgres=# DROP TABLE test_delay ;
DROP TABLE
```

[illegible]

```
postgres=# SELECT * FROM test_delay;
    id |     create_time
-------+---------------------
       1 | 2017-09-10 16:17:50
(1 row)
```

[illegible]test_delay[illegible]

[illegible] recovery_min_apply_delay[illegible]pg_wal[illegible]WAL[illegible]

[illegible]pg_wal[illegible]
[illegible]

12.6.3 recovery_min_apply_delay[illegible]

recovery_min_apply_delay[illegible]synchronous_commit[illegible]on[illegible]remote_apply[illegible]on[illegible]WAL[illegible]WAL[illegible]remote_apply[illegible]WAL[illegible]WAL[illegible]WAL[illegible]12.2.1[illegible]

[illegible]synchronous_commit[illegible]on[illegible]remote_apply[illegible]

[illegible]synchronous_commit[illegible]on[illegible]recovery_min_apply_delay[illegible]1[illegible]

synchronous_commit[illegible]pg_ctlreload[illegible]recovery_min_apply_delay[illegible]

[illegible]test_delay[illegible]

```
postgres=# INSERT INTO test_delay(id,create_time)
VALUES(1,now());
INSERT 0 1
```

[illegible]

```
postgres=# SELECT now(),create_time FROM test_delay;
-[ RECORD 1 ]-----------------------------
now         | 2017-09-10 16:58:22.526087+08
create_time | 2017-09-10 16:57:22
```

[illegible]synchronous_commit[illegible] on[illegible]

[illegible]synchronous_commit[illegible] remote_apply[illegible] recovery_min_apply_delay[illegible]1[illegible]

[illegible]SQL[illegible]test_delay[illegible]

```
postgres=# INSERT INTO test_delay(id,create_time)
VALUES(2,now());
--[illegible]
```

[illegible]SQL[illegible]SQL[illegible]

```
postgres=# \timing
Timing is on.
postgres=# INSERT INTO test_delay(id,create_time)
VALUES(3,now());
INSERT 0 1
Time: 60008.295 ms (01:00.008)
```

[illegible]SQL[illegible]60[illegible]INSERT[illegible]60[illegible]synchronous_commit[illegible]remote_apply[illegible]INSERT[illegible]INSERT[illegible]WAL[illegible]WAL[illegible]WAL[illegible]WAL[illegible]

[illegible]synchronous_commit[illegible]on[illegible]synchronous_commit[illegible]remote_apply[illegible]recovery_min_apply_delay[illegible]synchronous_commit[illegible]remote_apply[illegible]

12.7 [illegible]

[illegible]PostgreSQL[illegible]PostgreSQL9.6[illegible]PostgreSQL10[illegible]synchronous_standby_names[illegible]ANY[illegible]Quorum[illegible]PostgreSQL10[illegible]

[illegible]pghost3[illegible]12-4[illegible]

[illegible]12-4 [illegible]

主　　机	主　机　名	IP 地址	操 作 系 统	PostgreSQL 版本
主节点	pghost1	192.168.28.74	CentOS6.9	PostgreSQL10
备节点 1	pghost2	192.168.28.75	CentOS6.9	PostgreSQL10
备节点 2	pghost3	192.168.28.76	CentOS6.9	PostgreSQL10

[illegible]pg_stat_replication[illegible]

```
postgres=# SELECT
pid,usename,application_name,client_addr,state,sync_state,sy
nc_priority
           FROM pg_stat_replication ;
```

```
    pid  | usename | application_name | client_addr |
state   | sync_state | sync_priority
--------+--------+-------------+-----------+--------+---
-------+------------
 26030  | repuser | node2           |192.168.28.75 |
streaming | async     |       0
  2799  | repuser | node3           |192.168.28.76 |
streaming | async     |       0
(2 rows)
```

□□□□□□□□□node2□node3□□□□□□□sync_state□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□synchronous_standby_names□□□

12.7.1 synchronous_standby_names□□□□

synchronous_standby_names□string□□□□□□□□□□□□□□PostgreSQL10□□□□□□□□□□□□□□

·standby_name[□...]

·[FIRST]num_sync □standby_name[□...]□

·ANY num_sync□standby_name[□...]□

□□□□synchronous_standby_names=standby_name[□...]

standby_name□□□□□□□□□□□□□□□□□□$PGDATA/recovery.conf□□□□□□primary_conninfo□□application_name□□□9.5□□□9.5□□□□□□□□□□□□□□□□□

□□□□□'s1□s2'□□□s1□□□□□□s2□□□□□□□□□□s1□□□□□s2□□□□□□□□

□□□□synchronous_standby_names=[FIRST]num_sync□standby_name[□...]□

FIRST□□num_sync□□□□□□□□□□□□□□□□□

```
synchronous_standby_names =’FIRST 2(s1□s2□s3)’
```

□□□□□□□□□□□□□□□s1□s2□□□□□□□□s1□s2□□□□□□□□□□□□□□□□□□□□□□□□□□s1□s2□□□□□□WAL□□□□□□WAL□□□□□□□□□□□□□□□□s3□□□□□□□□□□s1□s2□□□□s3□□□□□□□□□□

□□□□ synchronous_standby_names=ANY num_sync□standby_name[□...]□

ANY□□□□quorum□□□□□□□□□□□□□□□□□□□num_sync□□□□□□□□□□□□□□□□s1□s2□s3□s4□□□□□□□□□□

```
synchronous_standby_names = 'ANY 2 (s1, s2, s3)'
```

ANY 2□□WAL□□□□□□WAL□□□□□□□□□□□□□□□s4□□□□□□□□s4□□□□□□

□□□□□□□□□□□□□□□□Quorum□□□□□□□□□□

12.7.2 □□□□□□□□□□

pghost1□□□□□postgresql.conf□□□□□□□□

```
synchronous_standby_names = 'first 1 (node2,node3)'
```

□□□□□□□first 1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□node2□□□□□node3□□□□□□□□□□□□□□□□□□□□

□□□□reload□□□□□□□□□□□

```
[postgres@pghost1 pg_root]$ pg_ctl reload
server signaled
```

□□□□□□□□□□□□□□□□□□□□□□□

```
postgres=# show synchronous_standby_names ;
    synchronous_standby_names
---------------------------
    first 1 (node2,node3)
(1 row)
```

□□□□□□pg_stat_replication□□□□□□□□

```
postgres=# SELECT
pid,usename,application_name,client_addr,state,sync_state,sy
nc_priority FROM pg_stat_replication ;
    pid | usename | application_name | client_addr |   state
| sync_state  |sync_priority
--------+--------+-------------+------------+----------+--
--------+-----------
    1536 | repuser | node2          | 192.168.28.75 |
streaming | sync       |           1
    2799 | repuser | node3          | 192.168.28.76 |
streaming | potential  |           2
(2 rows)
```

□□□□sync_state□□□□□□□node2□sync_state□□□□async□□□sync□□□node2□sync_priority□□□□1□node3□synce_state□□□

□async□□□□□□potential□□□node3□sync_priority□□□□2□□□□□sync□□□□□□□potential□□□□□□□□

□□□□node2□□□□□□□□□□□

```
[postgres@pghost2 pg_root]$ pg_ctl stop -m fast
waiting for server to shut down.... done
server stopped
```

□□□□□□□□pg_stat_replication□□□□□□□□

```
postgres=# SELECT
pid,usename,application_name,client_addr,state,sync_state,sy
nc_priority FROM pg_stat_replication ;
    pid  | usename | application_name | client_addr  |
state   | sync_state | sync_priority
--------+-------+--------------+-----------+--------+----
------+------------
    2799 | repuser | node3            | 192.168.28.76 |
streaming | sync       |          2
(1 row)
```

□□□□node3□sync_state□□□sync□□□□□□□□□

□□□□□□□□□□□□□□□□□□

```
postgres=# INSERT INTO test_delay(id,create_time) VALUES
(4,now());
INSERT 0 1
```

node2[illegible]node3[illegible]

```
[postgres@pghost3 pg_root]$ pg_ctl stop -m fast
waiting for server to shut down.... done
server stopped
```

[illegible]

```
postgres=# INSERT INTO test_delay(id,create_time) VALUES
(5,now());
--[illegible]
```

[illegible]INSERT[illegible]

[illegible]synchronous_standby_names='first 1[illegible]node2[illegible]node3[illegible]'[illegible]node2[illegible]node3[illegible]

12.7.3 [illegible]Quorum[illegible]

[illegible]Quorum[illegible]PostgreSQL10[illegible]synchronous_standby_names[illegible]ANY[illegible]

```
synchronous_standby_names = 'ANY 2 (node2,node3)'
```

[illegible]WAL[illegible]WAL[illegible]node2[illegible]node3[illegible]

[illegible]reload[illegible]

```
[postgres@pghost1 pg_root]$ pg_ctl reload
server signaled
```

[illegible]

```
postgres=# show synchronous_standby_names ;
    synchronous_standby_names
---------------------------
    ANY 2 (node2,node3)
(1 row)
```

[illegible]pg_stat_replication[illegible]

```
postgres=# SELECT
pid,usename,application_name,client_addr,state,sync_state,sy
nc_priority
          FROM pg_stat_replication ;
    pid  | usename|application_name| client_addr  | state
| sync_state | sync_priority
```

```
---------+-------+-----------+-----------+--------+--
--------+-----------
    26906 | repuser | node3          | 192.168.28.76 |
streaming| quorum    |           1
    26926 | repuser | node2          | 192.168.28.75 |
streaming| quorum    |           1
(2 rows)
```

[illegible]node2[illegible]node3[illegible]sync_state[illegible]quorum[illegible]sync_priority[illegible]1[illegible]Quorum[illegible]sync_priority[illegible]node2[illegible]

```
[postgres@pghost2 pg_root]$ pg_ctl stop
waiting for server to shut down.... done
server stopped
```

[illegible]pg_stat_replication[illegible]node3[illegible]

```
postgres=# SELECT
pid,usename,application_name,client_addr,state,sync_state,sy
nc_priority
          FROM pg_stat_replication ;
    pid | usename | application_name | client_addr | state |
sync_state | sync_priority
-------+-------+---------------+-----------+------+----
-------+------------
 26906 | repuser | node3           | 192.168.28.76 |
streaming | quorum|           1
(1 row)
```

□□□□□□□□□□□□□□□□□□□□

```
postgres=# INSERT INTO test_delay(id,create_time)
VALUES(5,now());
--□□□□□□□□
```

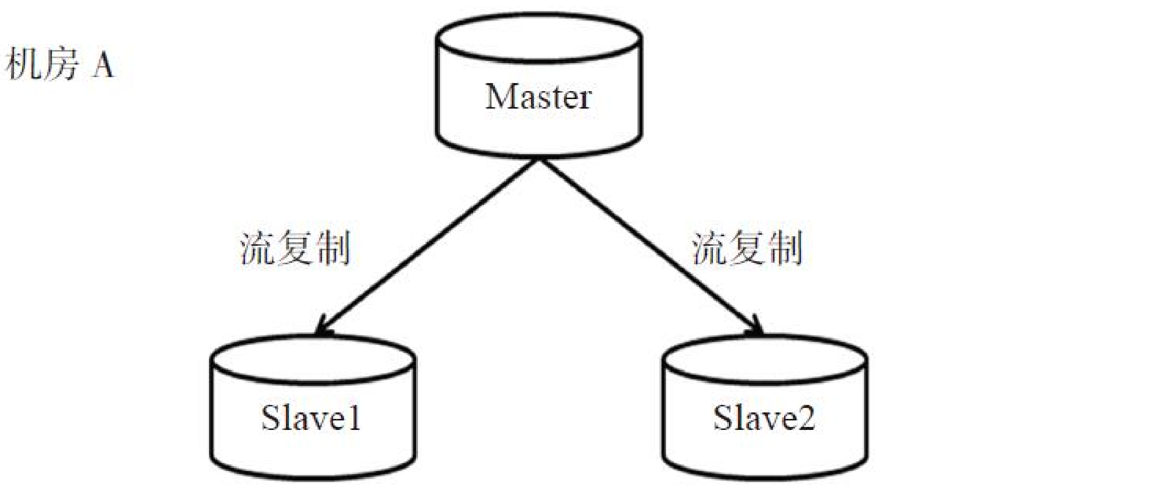
□□□□INSERT□□□□□□□□□□□□□□□□□□□□□□□□□'ANY 2□node2□node3□'□□□□□□□□□□□□□□□□□□□□□□□WAL□□□□□□WAL□□

12.8 [illegible]

[illegible] PostgreSQL[illegible]WAL[illegible]WAL[illegible]Cascading Replication[illegible]

12.8.1 [illegible]

[illegible]12-3[illegible]

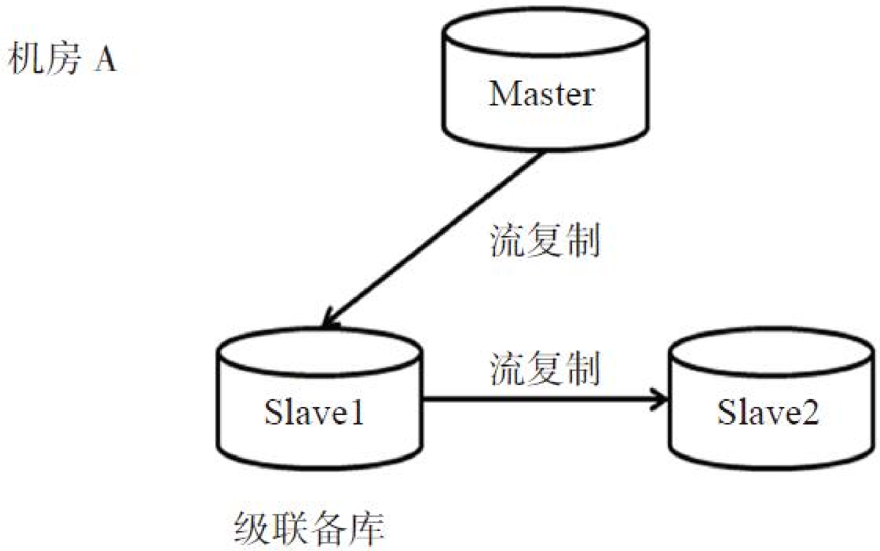


[illegible]12-3 [illegible]

[illegible]Master[illegible]Save1[illegible]Slave2[illegible] Save1[illegible]Slave2[illegible]Master[illegible]

□□□A□

□□□□□□□□□12-4□□□



□12-4 □□□□□□□□□□□□□

□□□□□□□□□□12-3□□□□□□□□□Slave2□□□□□□Master□□□□□□□Slave1□□□Slave1□□□□□□□□□Master□□□WAL□□□□□□□WAL□□□□□Slave2□□□□□□WAL□□□□□WAL□□□□□□□□□cascading standby□□□□Slave1□□

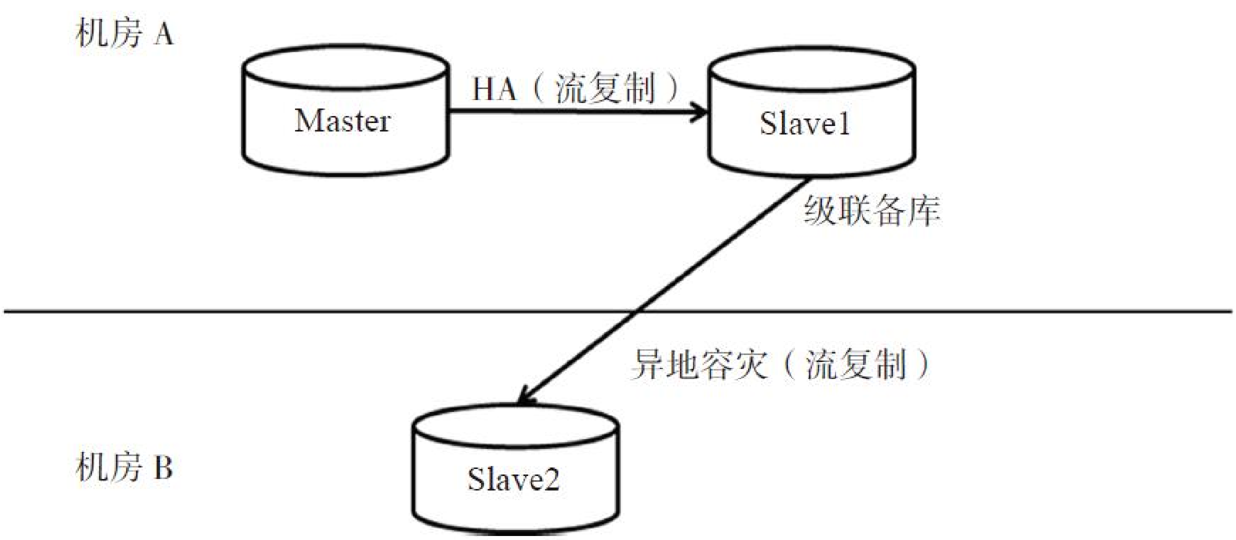
□□□□□□□□□□

·□□□□□□CPU□□□

·[illegible]

·[illegible]

[illegible]12-5[illegible]



[illegible]12-5 [illegible]+[illegible]

[illegible]Slave1[illegible]Master[illegible]14[illegible]PostgreSQL[illegible]Slave2[illegible]

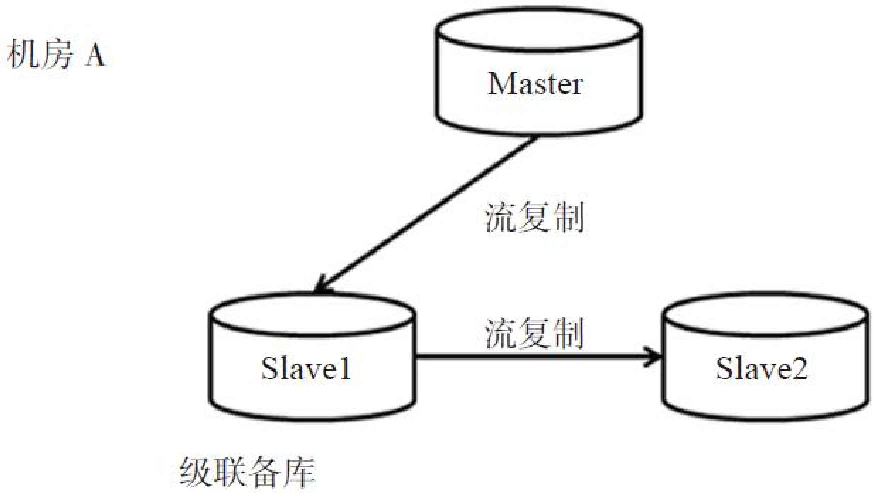
12.8.2 [illegible]

[illegible]12-5[illegible]

□12-5   □□□□□□□□

主　　机	主　机　名	IP 地址	操 作 系 统	PostgreSQL 版本
Master	pghost1	192.168.28.74	CentOS6.9	PostgreSQL10
Slave1	pghost2	192.168.28.75	CentOS6.9	PostgreSQL10
Slave2	pghost3	192.168.28.76	CentOS6.9	PostgreSQL10

□□□□□□□□12-6□



□12-6   □□□□□□□□□

□□□□Slave1□□□□□□□□Slave2□□□□□□□□Slave1□

□□□□Slave1□□□□□□□□□□□Slave1□recovery.conf□□□□□□□

```
recovery_target_timeline = 'latest'
standby_mode = on
primary_conninfo = 'host=192.168.28.74  port=1921
user=repuser application_name=slave1'
```

Slave1[illegible] Slave2[illegible]Slave2[illegible]

```
[postgres@pghost3 pg10]$ pg_basebackup -D
/database/pg10/pg_root -Fp -Xs -v -P -h 192.168.28.74 -p
1921 -U repuser
pg_basebackup: initiating base backup, waiting for
checkpoint to complete
pg_basebackup: checkpoint completed
pg_basebackup: write-ahead log start point: 4/4E000028 on
timeline 7
pg_basebackup: starting background WAL receiver
9424317/9424317 kB (100%), 3/3 tablespaces
pg_basebackup: write-ahead log end point: 4/4E004AA8
pg_basebackup: waiting for background process to finish
streaming ...
pg_basebackup: base backup completed
```

[illegible]Slave2[illegible]recovery.conf[illegible]

```
recovery_target_timeline = 'latest'
standby_mode = on
primary_conninfo = 'host=192.168.28.75  port=1921
user=repuser application_name=slave2'
```

[illegible]Slave2[illegible]

```
[postgres@pghost3 pg_root]$ pg_ctl start
```

□□Slave2□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□Master□□pg_stat_replication□□□□□□

```
postgres=# SELECT
pid,usename,application_name,client_addr,state,sync_state,sy
nc_priority
          FROM pg_stat_replication ;
    pid  |usename   | application_name   | client_addr  |
state  | sync_state | sync_priority
---------+---------+-----------------+------------+---------
-+------+----------
    25041 | repuser | slave1             | 192.168.28.75 |
streaming | async  |       0
(1 row)
```

□□□□□□□□□Master□Slave1□WAL□□□□

□Slave1□□pg_stat_replication□□□□□□

```
postgres=# SELECT
pid,usename,application_name,client_addr,state,sync_state,sy
nc_priority
          FROM pg_stat_replication ;
    pid  | usename | application_name | client_addr |  state
| sync_state | sync_priority
--------+--------+----------------+-----------+--------+----
-------+------------
    5002| repuser| slave2         |192.168.28.76|streaming|
async     |           0
(1 row)
```

□□□□□Slave1□Slave2□WAL□□□□□□□Slave1□□□□WAL□□□□□

□□□□□□□□□□Master□□□□□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ psql postgres postgres
postgres=# CREATE table t_sr6(id int4);
CREATE TABLE
postgres=# INSERT INTO t_sr6 VALUES (1);
INSERT 0 1
```

□Slave1□□□□□□□□□□□

```
[postgres@pghost2 pg_root]$ psql postgres postgres
postgres=# SELECT * FROM t_sr6;
 id
----
 1
(1 row)
```

Slave1□□□□t_sr6□□Slave2□□□□□□□□□□□

```
[postgres@pghost3 pg_root]$ psql postgres postgres
postgres=# SELECT * FROM t_sr6;
 id
----
 1
(1 row)
```

Slave2□□□□□□□

12.9 □□□□□□□□□

PostgreSQL□□9.0□□□□□□□□□□□□□□□□□□□□□□□□□“□□”□□□□□□□□□□□□□□□□□□□□□

12.9.1 □□□□□□□□□□□□□□□□□□

□□PostgreSQL□□

□□□□□□□□□□□□□□□□□□□□□□□□pghost1□□□pghost2□□□□□□□□

```
postgres=# SELECT
pid,usename,application_name,client_addr,state,sync_state,sy
nc_priority
               FROM pg_stat_replication ;
    pid  | usename | application_name | client_addr |
state   | sync_state | sync_priority
--------+---------+-----------------+-----------+---------
```

```
-+------------+-----------
    24924| repuser | node2          |192.168.28.75|
streaming | async     |         0
(1 row)
```

□□□□□□□□tbs_his□□□□□□□□pghost1□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ mkdir -p /database/pg10/pg_tbs/tbs_his
```

□□□□□□□□tbs_his□□□□□□□□□

```
postgres=# CREATE TABLESPACE tbs_his OWNER pguser LOCATION
'/database/pg10/pg_tbs/tbs_his';
CREATE TABLESPACE
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
2017-09-12 16:35:11.962 CST,,,16742,,59b79b9f.4166,6,,2017-
09-12 16:32:31 CST,
1/0,0,FATAL,58P01,"directory
""/database/pg10/pg_tbs/tbs_his"" does not exist",,"Create
this directory for the tablespace before restarting the
server.",,,"WAL redo at 3/AA17E8D0 for Tablespace/CREATE:
25227 ""/database/pg10/pg_tbs/tbs_his""",,,,""
2017-09-12 16:35:11.962 CST,,,16740,,59b79b9f.4164,3,,2017-
09-12 16:32:31 CST,,0,LOG,00000,"startup process (PID 16742)
exited with exit code 1",,,,,,,,,""
2017-09-12 16:35:11.963 CST,,,16740,,59b79b9f.4164,4,,2017-
09-12 16:32:31 CST,,0,LOG,00000,"terminating any other
active server processes",,,,,,,,,""
2017-09-12 16:35:11.971 CST,,,16740,,59b79b9f.4164,5,,2017-
```

```
09-12 16:32:31 CST,,0,LOG,00000,"database system is shut
down",,,,,,,,,""
```

[illegible]/database/pg10/pg_tbs/tbs_his[illegible]

```
[postgres@pghost2 ~]$ mkdir -p
/database/pg10/pg_tbs/tbs_his
```

[illegible]pghost2[illegible]

```
[postgres@pghost2 ~]$ pg_ctl start
```

[illegible]WAL[illegible]

```
postgres=# SELECT
pid,usename,application_name,client_addr,state,sync_state,sy
nc_priority
          FROM pg_stat_replication ;
    pid   | usename | application_name | client_addr  |
state   | sync_state | sync_priority
---------+---------+----------------+-------------+--------
--+----------+------------
    26841| repuser | node2           | 192.168.28.75 |
streaming | async     |              0
(1 row)
```

□□□□□□□□□□□□□□□

□□□

12.9.2 □□□□□□□□□□□□

□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□

```
ERROR:  canceling statement due to conflict with recovery
DETAIL:  User query might have needed to see row versions
that must be removed.
```

□□□PostgreSQL□MVCC□□□□□□□□□□□□□□□□□□□□□□□□□□□□autovacuum□□□□□□□□□□VACUUM□□□□□□□□□□□□□□□□□□□VACUUM□□

□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

·max_standby_streaming_delay[illegible]30[illegible]SQL[illegible]WAL[illegible]30[illegible]30[illegible]WAL[illegible]WAL[illegible]-1[illegible]WAL[illegible]WAL[illegible]

·hot_standby_feedback[illegible]on[illegible]off[illegible]

[illegible]pghost1[illegible]pghost2[illegible]postgresql.conf[illegible]

```
max_standby_streaming_delay = 10s
```

[illegible]max_standby_streaming_delay[illegible]10[illegible]reload[illegible]

```
[postgres@pghost2 pg_root]$ pg_ctl reload
server signaled
```

□□update_per2.sql□□□□□□

```
\set v_id random(1,1000000)
update test_per2 set flag='1' where id=:v_id;
```

pghost1□□pgbench□□□□□□□□□□□□120
□□□□□□□

```
pgbe nch -c 8 -T 120 -d postgres -U postgres -n N -M
prepared -f update_per2.sql> update_8.out 2>&1 &
```

□□□□□□□□□□□□□□□□□□□□□□□□

```
postgres=# \timing
Timing is on.
postgres=# SELECT pg_sleep(15),count(*) FROM test_per2;
ERROR: canceling statement due to conflict with recovery
DETAIL: User query might have needed to see row versions
that must be removed.
Time: 10433.102 ms (00:10.433)
```

□□□□□□□test_per2□□□□□□□□□□pg_sleep
□□□□□□□□10□□□□□□□□□□□□

□□□□□□□□□□□□□□

□□□□□□max_standby_streaming_delay □□□

□□□□□max_standby_streaming_delay□□□10□□□□□□□□□□□□□□□WAL□□□□□□□□□SQL□□□□□10□□□□□□□□□□□□□□□□□□□□□□□□-1□□□□□□□□□□□□□□□□60□□□

```
max_standby_streaming_delay = 60s
hot_standby_feedback = off
```

□□□hot_standby_feedback□□□□□off□□□□□□□□reload□□□□□□□□□□□

```
[postgres@pghost2 pg_root]$ pg_ctl reload
server signaled
```

□□□□□□pgbench□□□□□□□□□□□□□□□□□□

```
postgres=# SELECT pg_sleep(15),count(*) FROM test_per2;
    pg_sleep |  count
-------------+----------
             | 10000000
(1 row)

Time: 15327.394 ms (00:15.327)
```

□□□□□□□□15□□□□□□

[illegible]hot_standby_feedback[illegible]

hot_standby_feedback[illegible]on[illegible]

```
hot_standby_feedback = on
max_standby_streaming_delay = 10s
```

[illegible]hot_standby_feedback[illegible]on[illegible]max_standby_streaming_delay[illegible]10[illegible]reload[illegible]

```
[postgres@pghost2 pg_root]$ pg_ctl reload
server signaled
```

[illegible]pgbench[illegible]

```
postgres=# SELECT pg_sleep(15),count(*) FROM test_per2;
    pg_sleep |  count
-------------+----------
             | 10000000
(1 row)

Time: 15349.958 ms (00:15.350)
```

[illegible]15[illegible]

[illegible]max_standby_streaming_delay[illegible]-1[illegible]hot_standby_feedback[illegible]

12.9.3 [illegible]WAL[illegible]

[illegible]

```
FATAL,XX000,"could not receive data from WAL stream: ERROR:
requested WAL segment 000000010000000100000022 has already
been removed"
```

[illegible]WAL[illegible]000000010000000100000022[illegible]WAL[illegible]WAL[illegible]WAL[illegible]wal_keep_segment[illegible]WAL[illegible]WAL[illegible]

□□□□□□□□□□□□□□□□□□□□□1TB□□□□□□□□
□□□600GB□□□□□□□□□□□□□□□□□□□□□□□□□
□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□pghost1□□□□pghost2□□□□□□□postgresql.conf□□□□□□□

```
wal_keep_segments = 1
archive_mode = on
archive_command = 'cp %p /archive_dir/pg10/%f'
```

□wal_keep_segments□□□1□□□□pg_wal□□□□□□□WAL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□WAL□□□□□□/archive_dir/pg10□□□□□reload□□□□□□□□□□

```
[postgres@pghost1 pg_root]$ pg_ctl reload
server signaled
```

□□wal_keep_segments□□□□□1□□pg_wal□□□□□100□□WAL□□□□□□□□□□□□checkpoint□□□□□pg_wal□□□□WAL□□□□□□□□□□

```
postgres=# CHECKPOINT;
CHECKPOINT
```

[illegible]pg_wal[illegible]WAL[illegible]WAL[illegible]WAL[illegible]

[illegible]

```
[postgres@pghost2 ~]$ pg_ctl stop -m fast
waiting for server to shut down.... done
server stopped
```

[illegible]update[illegible]update_per2.sql[illegible]

```
\set v_id random(1,1000000)

update test_per2 set flag='1' where id=:v_id;
```

[illegible]pgbench[illegible]

```
pgbench -c 8 -T 120 -d postgres -U postgres -n N -M prepared
-f update_per2.sql >  update_8.out 2>&1 &
[1] 23803
```

[illegible]pgbench[illegible]pg_switch_wal[illegible]checkpoint[illegible]

□□□□□

```
postgres=# SELECT pg_switch_wal();
    pg_switch_wal
---------------
    4/32983D58
(1 row)
...□□□□

postgres=# checkpoint;
CHECKPOINT
...□□□□
```

□□□□□□□□□□pg_switch_wal□□□□□□□□WAL□□□□□□□□□□□□□□□□□□□□□□□

```
[postgres@pghost1 pg_root]$ ll /archive_dir/pg10/
total 352M
-rw------- 1 postgres postgres 16M Sep 14 21:34
000000070000000400000023
-rw------- 1 postgres postgres 16M Sep 14 21:34
000000070000000400000024
□□
```

□□WAL□□□□/archive_dir/pg10/□□□□□□□WAL□□□□□

□□□□□□□□□□□

```
[postgres@pghost2 ~]$ pg_ctl start
server started
```

[illegible]WAL[illegible]

```
2017-09-14 21:44:34.764 CST,,,25011,,59ba87c2.61b3,2,,2017-
09-14 21:44:34 CST,,0,FATAL,XX000,"could not receive data
from WAL stream: ERROR:  requested WAL segment
000000070000000400000023 has already been
removed",,,,,,,,,""
```

[illegible]000000070000000400000023[illegible]$PGDATA/pg_wal[illegible]000000070000000400000023[illegible]

```
[postgres@pghost1 pg_root]$ ll
pg_wal/000000070000000400000023
ls: cannot access pg_wal/000000070000000400000023: No such
file or directory
```

[illegible]/archive_dir/pg10/[illegible]WAL[illegible]

```
[postgres@pghost1 pg_root]$ ll
/archive_dir/pg10/000000070000000400000023
-rw------- 1 postgres postgres 16M Sep 14 21:34
/archive_dir/pg10/000000070000000400 000023
```

□□□□□□□□□□□□□□WAL□□□□□□□□□□□□□□□□□□□□□

```
[postgres@pghost1 pg_root]$ scp /archive_dir/pg10/*
postgres@pghost2:/archive_dir/pg10
postgres@pghost2's password:
```

□□□□□□□□recovery.conf□□□□□□□□□□□□□

```
restore_command = 'cp /archive_dir/pg10/%f %p'
```

restore_command□□□□□□shell□□□□□□□□□□WAL□□□□□WAL□□□□□□□□□□□□□□□□

```
[postgres@pghost2 pg_root]$ pg_ctl restart
```

□□□□□□□□□□□□□□□

```
2017-09-14 21:55:16.904 CST,,,25224,,59ba8a44.6288,1,,2017-
09-14 21:55:16 CST,,0,LOG,00000,"ending log output to
stderr",,"Future log output will go to log destination
""csvlog"".",,,,,,,,""
2017-09-14 21:55:16.911 CST,,,25226,,59ba8a44.628a,1,,2017-
09-14 21:55:16 CST,,0,LOG,00000,"database system was shut
down in recovery at 2017-09-14 21:55:11 CST",,,,,,,,,""
2017-09-14 21:55:16.913 CST,,,25226,,59ba8a44.628a,2,,2017-
09-14 21:55:16 CST,,0,LOG,00000,"entering standby
mode",,,,,,,,,""
2017-09-14 21:55:16.933 CST,,,25226,,59ba8a44.628a,3,,2017-
09-14 21:55:16 CST,1/0,0,LOG,00000,"redo starts at
```

```
4/1A00D8A8",,,,,,,,,""
2017-09-14 21:55:18.052 CST,,,25226,,59ba8a44.628a,4,,2017-
09-14 21:55:16 CST,1/0,0,LOG,00000,"restored log file
""000000070000000400000023"" from archive",,,,,,,,,""
2017-09-14 21:55:18.526 CST,,,25226,,59ba8a44.628a,5,,2017-
09-14 21:55:16 CST,1/0,0,LOG,00000,"consistent recovery
state reached at 4/23D6EA98",,,,,,,,,""
2017-09-14 21:55:18.527 CST,,,25224,,59ba8a44.6288,2,,2017-
09-14 21:55:16 CST,,0,LOG,00000,"database system is ready to
accept read only connections",,,,,,,,,""
2017-09-14 21:55:18.552 CST,,,25226,,59ba8a44.628a,6,,2017-
09-14 21:55:16 CST,1/0,0,LOG,00000,"restored log file
""000000070000000400000024"" from archive",,,,,,,,,""
2017-09-14 21:55:18.685 CST,,,25226,,59ba8a44.628a,7,,2017-
09-14 21:55:16 CST,1/0,0,LOG,00000,"restored log file
""000000070000000400000025"" from archive",,,,,,,,,""
...
2017-09-14 21:55:22.192 CST,,,25264,,59ba8a4a.62b0,1,,2017-
09-14 21:55:22 CST,,0,LOG,00000,"started streaming WAL from
primary at 4/38000000 on timeline 7",,,,,,,,,""
```

[illegible]WAL[illegible]000000070000000400000023[illegible]WAL[illegible]“started streaming WAL”[illegible]

[illegible]WAL[illegible]WAL[illegible]

[illegible]

[illegible]wal_keep_segments[illegible]$PGDATA/pg_wal[illegible]WAL[illegible]

[illegible]WAL[illegible]pg_wal[illegible]

[illegible]WAL[illegible]WAL[illegible]WAL[illegible]

[illegible]Replication slots[illegible]“[illegible]”[illegible]WAL[illegible]WAL[illegible]pg_wal[illegible]pg_wal[illegible]

[illegible]postgresql.conf[illegible]

```
max_replication_slots = 1
wal_keep_segments = 1
```

[illegible]

[illegible]

```
postgres=# SELECT * FROM
pg_create_physical_replication_slot('phy_slot1');
    slot_name | lsn
-------------+-----
    phy_slot1 |
(1 row)
```

[illegible]pg_replication_slots[illegible]pg_replication_slots[illegible]

```
postgres=# SELECT
slot_name,plugin,slot_type,database,active,active_pid,xmin
           FROM pg_replication_slots ;
    slot_name | plugin | slot_type | database | active |
active_pid | xmin
-------------+--------+----------+---------+--------+----
--------+------
    phy_slot1 |        | physical  |          | f      |
|
(1 row)
```

·slot_name[illegible]

·plugin[illegible]

·slot_type[illegible]physical[illegible]logical[illegible]

·database[illegible]

·active[illegible]t[illegible]

·active_pid[illegible]

·xmin[illegible]

[illegible]recovery.conf[illegible]primary_slot_name[illegible]

```
recovery_target_timeline = 'latest'
standby_mode = on
primary_conninfo = 'host=192.168.28.74  port=1921
user=repuser application_name=slave1'
primary_slot_name = 'phy_slot1'
```

[illegible]

```
[postgres@pghost2 pg_root]$ pg_ctl restart -m fast
waiting for server to shut down.... done
server stopped
```

[illegible]pg_replication_slots[illegible]

```
slot_name,plugin,slot_type,database,active,active_pid,xmin
           FROM pg_replication_slots ;
    slot_name | plugin | slot_type | database | active |
active_pid | xmin
--------------+--------+-----------+---------+--------+----
--------+------
    phy_slot1 |        | physical  |          | t      |
23833 |
(1 row)
```

□□active□□□□□t□□□□□□□23833□□□□□□□□□□□

```
[postgres@pghost1 pg_root]$ ps -ef | grep 23833
postgres 23833 19503  0 14:40 ?        00:00:00 postgres:
wal sender process repuser 192.168.28.75(33141) streaming
4/52044950
```

23833□□□□□□□WAL□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□WAL□□□□□□□

□□□□□□□□□□

```
[postgres@pghost2 ~]$ pg_ctl stop -m fast
waiting for server to shut down.... done
server stopped
```

□□□□□□□□□□□□□□□pgbench□□□□□□□□□□□□

```
pgbench -c 8 -T 120 -d postgres -U postgres -n N -M prepared
-f update_per2.sql >  update_8.out 2>&1 &
```

□□□□□□□□□□□□□□□□□□pg_switch_wal□□□checkpoint□□□

[illegible]

```
[postgres@pghost2 ~]$ pg_ctl start
```

[illegible]WAL[illegible]$PGDATA/pg_wal[illegible]WAL[illegible]

12.10 [illegible]

PostgreSQL10[illegible]8[illegible]PostgreSQL10[illegible]Logical Replication[illegible]2ndquadrant[illegible]

[illegible]

·[illegible]

·[illegible]

·[illegible]PostgreSQL[illegible]

·[illegible]PostgreSQL[illegible]

[illegible]WAL[illegible]WAL[illegible]WAL[illegible]WAL[illegible]logical decoding[illegible]WAL[illegible]WAL[illegible]WAL[illegible]WAL[illegible]

12.10.1 □□□□

□□□□logical decoding□□□□□□□□□□□□□□□□□□□□□□□□□□□□□WAL□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□wal_level□□logical□□□max_replication_slots□□□□1□□□□□

```
wal_level = logical
max_replication_slots = 8
```

wal_level□□□WAL□□□□□□□minimal□replica□logical□□□□12.1.1□□□□□□□□□□□□□□□□□□□□□□□

max_replication_slots□□□□□□□□□□□□□□□□□□□□□□

pghost1□□□□□□□□□□□□□

```
postgres=# SELECT
pg_create_logical_replication_slot('logical_slot1','test_dec
oding');
    slot_name     |    lsn
------------------+------------
    logical_slot1 | 4/85004210
(1 row)
□□pg_replication_slots□□□□□□□□
postgres=# SELECT
```

```
slot_name,plugin,slot_type,database,active,restart_lsn
               FROM pg_replication_slots;
    slot_name     |    plugin     | slot_type | database |
active | restart_lsn
------------------+---------------+-----------+----------+--
------+-------------
    logical_slot1 | test_decoding | logical   | postgres | f
| 4/850041D8
    phy_slot1     |               | physical  |          | t
| 4/85016670
(2 rows)
```

[illegible]logical_slot1[illegible]pgh_slot1[illegible]12.9.3[illegible]WAL[illegible]

[illegible]logical_slot1[illegible]

```
postgres=# SELECT * FROM
pg_logical_slot_get_changes('logical_slot1',null,null);
    lsn | xid | data
--------+-----+------
(0 rows)
```

pg_logical_slot_get_changes[illegible]

[illegible]DDL[illegible]

```
postgres=# CREATE TABLE t_logical(id int4);
CREATE TABLE
postgres=# SELECT * FROM
pg_logical_slot_get_changes('logical_slot1',null,null);
    lsn        |   xid    |      data
---------------+----------+-----------------
    4/85094B38 | 42976847 | BEGIN 42976847
    4/850A9DE8 | 42976847 | COMMIT 42976847
(2 rows)
```

□□□□□□□□□□□□□□□DDL□□□□□□□□□□□DDL□□□□□DDL□□□□□□□□□

□□□□logical_slot1□□□□□□□□□□□□□□□□

```
postgres=# SELECT * FROM
pg_logical_slot_get_changes('logical_slot1',null,null);
    lsn | xid | data
--------+-----+------
(0 rows)
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□t_logical□□□□□□□□□□□logical_slot1□□□□□□□□□□□□

```
postgres=# INSERT INTO t_logical VALUES (1);
INSERT 0 1
postgres=# SELECT * FROM
pg_logical_slot_get_changes('logical_slot1',null,null);
        lsn    |   xid    |                     data
---------------+----------+---------------------------------
--------------
```

```
    4/850AB898 | 42976848 | BEGIN 42976848
    4/850AB898 | 42976848 | table public.t_logical: INSERT:
id[integer]:1
    4/850AB908 | 42976848 | COMMIT 42976848
(3 rows)
```

[illegible]INSERT[illegible]

[illegible] pg_logical_slot_get_changes[illegible]
pg_logical_slot_peek_changes[illegible]
pg_logical_slot_get_changes[illegible]pg_logical_slot_get_changes[illegible]
pg_logical_slot_peek_changes[illegible]

[illegible]
pg_recvlogical[illegible]

```
postgres=# INSERT INTO t_logical VALUES (3);
INSERT 0 1
```

[illegible]pg_recvlogical[illegible]
logical_slot1[illegible]

```
[postgres@pghost1 ~]$ pg_recvlogical -d postgres --slot
logical_slot1 --start -f -
BEGIN 42976850
table public.t_logical: INSERT: id[integer]:3
COMMIT 42976850
□□
```

-d□□□□□□□□--slot□□□□□□□□□□--start□□□□--slot□□□□□□□□□□□□□□□□□□-f□□□□□□□□□□□□□□□“-”□□□□□□□□□□□□□□□□□□□INSERT□□□□□□□□

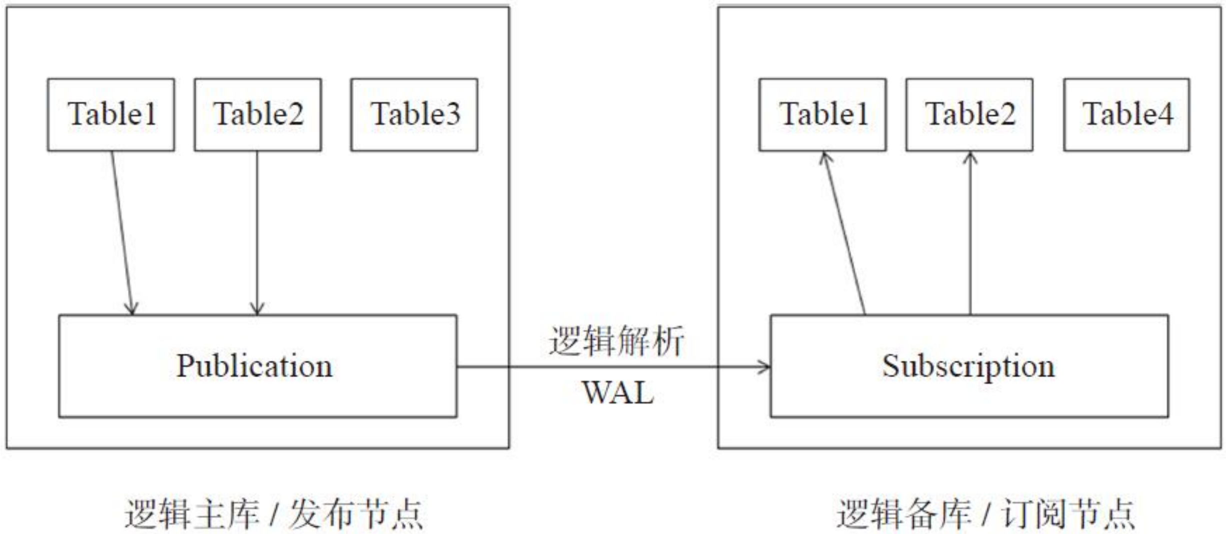
□□□□□□□□□□□□□□□□□□□□

```
postgres=# SELECT pg_drop_replication_slot('logical_slot1');
    pg_drop_replication_slot
--------------------------

(1 row)
```

12.10.2 □□□□

□□□□□□12-7□□



[illegible]12-7 [illegible]

[illegible]PostgreSQL[illegible]table1[illegible]table2[illegible]Publication[illegible]Subscription[illegible]table1[illegible]table2[illegible]

[illegible]Publication[illegible]WAL[illegible]Subscription[illegible]WAL[illegible]

[illegible]Publication[illegible]Subscription[illegible]

Publication[illegible]PostgreSQL[illegible]Publication[illegible]

[illegible]replica identity[illegible]DELETE/UPDAE[illegible]full[illegible]UPDATE/DELETE[illegible]

Subscription[illegible]Subscription[illegible]DDL[illegible]DDL[illegible]WAL[illegible]

12.10.3 [illegible]

[illegible]12-6[illegible]

[illegible]12-6 [illegible]

角色	主机名	IP	端口	库名	用户名	版本
发布节点	pghost1	192.168.28.74	1921	mydb	pguser	PostgreSQL10
订阅节点	pghost3	192.168.28.76	1923	des	pguser	PostgreSQL10

[illegible]postgresql.conf[illegible]

```
wal_level = logical
max_replication_slots = 8
```

```
max_wal_senders = 10
```

·wal_level[illegible]logical[illegible]

·max_replication_slots[illegible]

·max_wal_senders[illegible]WAL[illegible]max_replication_slots[illegible]

[illegible]postgresql.conf[illegible]

```
max_replication_slots = 8
max_logical_replication_workers = 8        # taken from
max_worker_processes
```

·max_replication_slots[illegible]

·max_logical_replication_workers[illegible]4[illegible]

[illegible]max_logical_replication_workers[illegible]max_worker_processes[illegible]max_worker_processes[illegible]

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
REPLICATION□□□□□□□□□□

```
CREATE USER logical_user
    REPLICATION
    LOGIN
    CONNECTION LIMIT 8
    ENCRYPTED PASSWORD 'logical_user';
```

□□□□□□□□□□REPLICATION□□□□□□□□□□□□□
SUPERUSER□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
logical_user□□□□□logical_user□□□□□□□□□□□□□□□
□□□□□□□□

□□□□□□□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ psql mydb pguser
psql (10.0)
Type "help" for help.

mydb=> CREATE TABLE t_lr1(id int4,name text);
CREATE TABLE
mydb=> INSERT INTO t_lr1 VALUES (1,'a');
INSERT 0 1
```

□□□□□□□□□t_lr1□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□

```
mydb=> CREATE PUBLICATION pub1 FOR TABLE t_lr1;
CREATE PUBLICATION
```

[illegible]

```
CREATE PUBLICATION name
    [ FOR TABLE [ ONLY ] table_name [ * ] [, ...]
        | FOR ALL TABLES ]
    [ WITH ( publication_parameter [= value] [, ... ] ) ]
```

·name[illegible]

·FOR TABLE[illegible]

·FOR ALL TABLES[illegible]PostgreSQL[illegible]PostgreSQL[illegible]

[illegible]pg_publication[illegible]

```
mydb=> SELECT * FROM pg_publication;
    pubname | pubowner | puballtables | pubinsert |
pubupdate | pubdelete
------------+----------+--------------+-----------+---------
--+-----------
    pub1    |    16391 | f            | t         | t
| t
(1 row)
```

·pubname□□□□□□□□

·pubowner□□□□□□□□□pg_user□□□usesysid□□□□□

·puballtables□□□□□□□□□□□□□□t□□□□□□□□□□□□□□□□□□□□□□□

·pubinsert□t□□□□□□□□INSERT□□□

·pubupdate□t□□□□□□□□UPDATE□□□

·pubdelete□t□□□□□□□□DELETE□□□

□□□□□□□□t_lr1□□□□□□□□□□□□□□□□□□□□□

```
[des@pghost3 ~]$ psql des pguser
psql (10.0)
Type "help" for help.

des=> CREATE TABLE t_lr1(id int4,name text);
CREATE TABLE
```

□□□□□□□□□□□□□□□□□□□□

```
CREATE SUBSCRIPTION subscription_name
    CONNECTION 'conninfo'
    PUBLICATION publication_name [, ...]
    [ WITH ( subscription_parameter [= value] [, ... ] ) ]
```

·subscription_name□□□□□□□□

·CONNECTION□□□□□□□□□□□□□□□host□port□dbname□user□password□□□□□□□□□□□□□□□□□□□□□□~/.pgpass□□□□□

·PUBLICATION□□□□□□□□□□□□□

·WITH□subscription_parameter[=value][□...]□□□□□□□□□□copy_data□boolean□□create_slot□boolean□□enabled□boolean□□slot_name□string□□□□□□□□□□□□

□□□□□□□□□□□□□□□□~/.pgpass□□□□□□□□□□□

```
192.168.28.74:1921:mydb:logical_user:logical_user
```

□~/.pgpass□□□□□□□□□□□□□□

```
[des@pghost3 ~]$ chmod 0600 .pgpass
```

□□□□□□□pg_hba.conf□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
[des@pghost3 ~]$ psql des postgres
psql (10.0)
Type "help" for help.

des=# CREATE SUBSCRIPTION sub1 CONNECTION
'host=192.168.28.74 port=1921
        dbname=mydb user=logical_user' PUBLICATION pub1;
NOTICE:  created replication slot "sub1" on publisher
CREATE SUBSCRIPTION
```

[illegible]sub1[illegible]

```
mydb=> SELECT
slot_name,plugin,slot_type,database,active,restart_lsn
        FROM pg_replication_slots where slot_name='sub1';
    slot_name |  plugin  | slot_type | database | active |
restart_lsn
--------------+----------+-----------+----------+--------+--
-----------
    sub1      | pgoutput | logical   | mydb     | t      |
4/9793A6F0
(1 row)
```

[illegible]plugin[illegible]pgoutput[illegible]plugin[illegible]

[illegible]pg_subscription[illegible]

```
des=# SELECT * FROM pg_subscription;
-[ RECORD 1 ]---+-------------------------------------------
-------------
```

```
subdbid         | 16387
subname         | sub1
subowner        | 10
subenabled      | t
subconninfo     | host=192.168.28.74 port=1921 dbname=mydb
user=logical_user
subslotname     | sub1
subsynccommit   | off
subpublications | {pub1}
```

·subdbid[illegible]OID[illegible]pg_database.oid[illegible]

·subname[illegible]

·subowner[illegible]

·subenabled[illegible]

·subconninfo[illegible]

·subslotname[illegible]

·subpublications[illegible]

[illegible]t_lr1[illegible]

```
des=> SELECT * FROM t_lr1;
 id | name
----+------
(0 rows)
```

□□□□□□t_lr1□□□□□□□□□□□□□□□□□□□□□□
□□

```
2017-10-01 14:39:45.795 CST,,,16650,,59d08db1.410a,1,,2017-
10-01 14:39:45 CST,4/47,0,LOG,00000,"logical replication
table synchronization worker for subscription ""sub1"",
table ""t_lr1"" has started",,,,,,,,,""
2017-10-01 14:39:45.875 CST,,,16650,,59d08db1.410a,2,,2017-
10-01 14:39:45 CST,4/50,0,ERROR,XX000,"could not start
initial contents copy for table ""pguser.t_lr1"": ERROR:
permission denied for schema pguser",,,,,,,,,""
2017-10-01 14:39:45.875 CST,,,6054,,59cf4455.17a6,286,,2017-
09-30 15:14:29 CST,,0,LOG,00000,"worker process: logical
replication worker for subscription 16396 sync 16389 (PID
16650) exited with exit code 1",,,,,,,,,""
```

□□□□□□□□□□□□□□□□sub1□□t_lr1□□□□□□□
□□□□□□□□□□□□□□□□□□□□□pguser□□□□□□□

□□□□□□□logical_user□□□□□□□□
logical_user□□□□□□□□

```
mydb=> GRANT USAGE ON SCHEMA pguser TO logical_user;
GRANT
mydb=> GRANT SELECT ON t_lr1 TO logical_user;
GRANT
```

□□□pguser□□□□□□□□□logical_user□□□□□
□□t_lr1□SELECT□□□□logical_user□□□

□□□□□□□□□□□□□□□□□□□

```
2017-10-01 14:45:18.893 CST,,,16755,,59d08efe.4173,1,,2017-
10-01 14:45:18 CST,4/400,0,LOG,00000,"logical replication
table synchronization worker for subscription ""sub1"",
table ""t_lr1"" has started",,,,,,,,,""
2017-10-01 14:45:18.951 CST,,,16755,,59d08efe.4173,2,,2017-
10-01 14:45:18 CST,4/404,0,LOG,00000,"logical replication
table synchronization worker for subscription ""sub1"",
table ""t_lr1"" has finished",,,,,,,,,""
```

[illegible]sub1[illegible]t_lr1[illegible]

```
des=> SELECT * from t_lr1 ;
    id | name
-------+------
    1  | a
(1 row)
```

[illegible]t_lr1[illegible]

[illegible]WAL[illegible]

```
postgres: wal sender process logical_user
192.168.28.76(47464) idle
```

[illegible]WAL[illegible]

```
postgres: bgworker: logical replication worker for
subscription 16396
```

[illegible]DML[illegible]

12.10.4 [illegible]DML[illegible]

[illegible]INSERT/UPDATE/DELTE[illegible]

[illegible]

```
mydb=> INSERT INTO t_lr1 VALUES (2,'b');
INSERT 0 1
```

[illegible]

```
des=> SELECT * FROM t_lr1 WHERE id=2;
    id | name
-------+------
    2  | b
(1 row)
```

[illegible]

```
mydb=> UPDATE t_lr1 SET name='bb' WHERE id=2;
ERROR:  cannot update table "t_lr1" because it does not have
replica identity and publishes updates
```

```
HINT:  To enable updating the table, set REPLICA IDENTITY
using ALTER TABLE.
```

[illegible]t_lr1[illegible]replica identity[illegible]12.10.2[illegible]UPDATE/DELETE[illegible]replica identity[illegible]t_lr1[illegible]

[illegible]t_lr1[illegible]

```
mydb=> ALTER TABLE t_lr1 ADD PRIMARY KEY(id);
ALTER TABLE
```

[illegible]t_lr1[illegible]

```
des=> ALTER TABLE t_lr1 ADD PRIMARY KEY(id);
ALTER TABLE
```

[illegible]

[illegible]

```
mydb=> UPDATE t_lr1 SET name='bb' WHERE id=2;
UPDATE 1
```

[illegible]t_lr1[illegible]UPDATE[illegible]

[illegible]

```
des=> SELECT * FROM t_lr1 WHERE id=2;
    id | name
-------+------
    2  | bb
(1 row)
```

[illegible]UPDATE[illegible]

[illegible]DELETE[illegible]ID[illegible]2[illegible]

```
mydb=> DELETE FROM t_lr1 WHERE id=2;
DELETE 1
```

[illegible]

```
des=> SELECT * FROM t_lr1 WHERE id=2;
    id | name
-------+------
(0 rows)
```

[illegible]ID[illegible]2[illegible]DELETE[illegible]

[illegible]t_lr1[illegible]t_lr1[illegible]t_lr1[illegible]

12.10.5 [illegible]

[illegible]

[illegible]t_big[illegible]1000[illegible]

```
mydb=> CREATE TABLE t_big(id int4 primary key,
create_time timestamp(0) without time zone default
clock_timestamp(),
name character varying(32));)
CREATE TABLE

mydb=> INSERT INTO t_big(id,name)
SELECT n,n*random()*10000 FROM generate_series(1,10000000)
n;
INSERT 0 10000000
```

[illegible]t_big[illegible]SELECT[illegible]logical_user[illegible]

```
mydb=> GRANT SELECT ON t_big TO logical_user;
GRANT
```

[illegible]t_big[illegible]pub1[illegible]

```
mydb=> ALTER PUBLICATION pub1 ADD TABLE t_big;
ALTER PUBLICATION
```

[illegible]\dRp+[illegible]

```
mydb=> \dRp+ pub1
            Publication pub1
    Owner  | All tables | Inserts | Updates | Deletes
-----------+------------+---------+---------+---------
    pguser | f          | t       | t       | t
Tables:
        "pguser.t_big"
        "pguser.t_lr1"
```

[illegible]t_big[illegible]pub1[illegible]

[illegible]pg_publication_tables[illegible]

```
mydb=> SELECT * FROM pg_publication_tables ;
    pubname | schemaname | tablename
------------+------------+-----------
    pub1    | pguser     | t_lr1
    pub1    | pguser     | t_big
(2 rows)
```

[illegible]t_big[illegible]

```
des=> CREATE TABLE t_big(id int4 primary key,
create_time timestamp(0) without time zone default
clock_timestamp(),
name character varying(32));)
CREATE TABLE
```

[illegible]t_big[illegible]t_big[illegible]

```
des=# ALTER SUBSCRIPTION sub1 REFRESH PUBLICATION ;
ALTER SUBSCRIPTION
```

[illegible]t_big[illegible]COPY[illegible]33[illegible]t_big[illegible]

```
des=# SELECT COUNT(*) FROM pguser.t_big;
    count
----------
    10000000
(1 row)
```

[illegible]t_big[illegible]

[illegible]t_big[illegible]t_big[illegible]pub1[illegible]

```
mydb=> ALTER PUBLICATION pub1 DROP TABLE t_big;
ALTER PUBLICATION
```

[illegible]t_big[illegible]

12.10.6 [illegible]

[illegible]

[illegible]sub1[illegible]

```
des=# ALTER SUBSCRIPTION sub1 DISABLE ;
ALTER SUBSCRIPTION
```

[illegible]pg_subscription[illegible]subenabled[illegible]

```
des=# SELECT subname,subenabled,subpublications FROM
pg_subscription;
    subname | subenabled | subpublications
------------+------------+-----------------
    sub1    | f          | {pub1}
(1 row)
```

[illegible]WAL[illegible]WAL[illegible]t_lr1[illegible]

[illegible]

```
des=# ALTER SUBSCRIPTION sub1 ENABLE ;
ALTER SUBSCRIPTION
```

□□pg_subscription□□□subenabled□□□□□□t□□□□□□

```
des=# SELECT subname,subenabled,subpublications FROM
pg_subscription;
    subname | subenabled | subpublications
------------+------------+----------------
    sub1    | t          | {pub1}
(1 row)
```

12.10.7 □□□□□□□□□□□□□

□□

□□□□□□□□□□□□□

·□□□□□wal_level□□□□□□□logical□

·□□□□□□□□□□□□□□□replication□□□□□

·□□□□□□□□□□□□□□□□UPDATE/DELETE□□□□□□□□□□□□□□□□□□□□□□□

·[illegible]INSERT[illegible]UPDATE[illegible]DELETE DML[illegible]

·[illegible]

·[illegible]

·[illegible]

·[illegible]pg_hba.conf[illegible]

·[illegible]DDL[illegible]DDL[illegible]DDL[illegible]

[illegible]

·[illegible]

·[illegible]

·[illegible]

·[illegible]

·[illegible]

·□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

·□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□UPDATE/DELETE□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

·DDL□□□□□□□□□□□□□□□□□□DDL□□□□DDL□□□□□□□□□□□□□□□□□□□□□□□□□□DDL□□□

·□□□

·TRUNCATE□□□□□□□□

·□□□□Large Object□□□□□□□□□

□□□□PostgreSQL10□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

12.10.8 □□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□INSERT□□□□□□□□□□□□□□□□UPDATE□□□□□

□□□□□□INSERT□□□□□

□□□□□□□□□□t_per1□□□□□SELECT□□□□□□□□□□logical_user□□□□□□

```
mydb=> CREATE TABLE t_per1 (id int4,name text,
create_time timestamp(0) without time zone DEFAULT '2000-01-
01 00:00:00');
CREATE TABLE

mydb=> GRANT SELECT ON t_per1 TO logical_user;
GRANT
```

□□□□□□□□pub1□□□t_per1□□□□□□

```
mydb=> ALTER PUBLICATION pub1 ADD TABLE t_per1 ;
ALTER PUBLICATION
```

□□□□□□□□□□□□□□□□

```
des=> CREATE TABLE t_per1 (id int4,name text,
create_time timestamp(0) without time zone DEFAULT '2000-01-
01 00:00:00');
CREATE TABLE
```

□□□□□□□□□sub1□□□□□

```
des=# ALTER SUBSCRIPTION sub1 REFRESH PUBLICATION ;
ALTER SUBSCRIPTION
```

□□□□□□□insert_t.sql□□□□□□□□□□□□

```
\set v_id random(1,1000000)

INSERT INTO t_per1(id,name) VALUES (:v_id,:v_id||'a');
```

□□□□pgbench□□INSERT□□□□□□□□□□

```
pgbench -c 8 -T 120 -d mydb -U pguser -n N -M prepared -f
insert_t.sql > insert_t.out 2>&1 &
```

□INSERT□□□□□□□□□□□□□□□□□□SQL□□□□□□

```
mydb=# SELECT pid,usename,state,
        pg_wal_lsn_diff(pg_current_wal_lsn(),replay_lsn)
replay_dely
        FROM pg_stat_replication WHERE
application_name='sub1';
    pid |   usename    |   state   | replay_dely
--------+-------------+-----------+-------------
    457 | logical_user | streaming |        11936
(1 row)
```

replay_dely□□□□□□WAL□□□□□□□□□□□□□□□□□SQL□WAL□□□□□□□11MB□□□□□□□□

□□□□□□UPDATE□□□□□

□□□□□□□□t_per2□□□□1000□□□□□□□□□

```
mydb=> CREATE TABLE t_per2 (id int4 primary key,
name text,
create_time timestamp(0) without time zone DEFAULT '2000-01-
01 00:00:00');
CREATE TABLE
mydb=> INSERT INTO t_per2(id) SELECT
generate_series(1,10000000);
INSERT 0 10000000
```

□□□□□□□□□□□□□□□□

```
mydb=> GRANT SELECT ON t_per2 TO logical_user ;
GRANT
mydb=> ALTER PUBLICATION pub1 ADD TABLE t_per2;
ALTER PUBLICATION
```

□□□□□□□t_per2□□□□□□□□□□□□

```
des=> CREATE TABLE t_per2 (id int4 primary key,
name text,
create_time timestamp(0) without time zone DEFAULT '2000-01-
01 00:00:00');
CREATE TABLE

des=# ALTER SUBSCRIPTION sub1 REFRESH PUBLICATION ;
ALTER SUBSCRIPTION
```

□□40□□□□□□□□t_per2□□□□□□□□□

□□□□□□□update_t.sql□□□□□□□□□□

```
\set v_id random(1,1000000)

update t_per2 set create_time=clock_timestamp() where
id=:v_id;
```

□□□□pgbench□□UPDATE□□□□□□□□□□

```
pgbench -c 8 -T 120 -d mydb -U pguser -n N -M prepared -f
update_t.sql > update_t.out 2>&1 &
```

UPDATE□□□□□□□□□□□□□□□SQL□□□□□□

```
mydb=# SELECT pid,usename,state,
            pg_wal_lsn_diff(pg_current_wal_lsn(),replay_lsn)
replay_dely
        FROM pg_stat_replication WHERE
application_name='sub1';
    pid |   usename    |   state   | replay_dely
--------+--------------+-----------+-------------
    457 | logical_user | streaming | 12352
```

□□□□□□SQL□replay_dely□□□12MB□□□□□
□□□

12.11 [illegible]

[illegible]PostgreSQL[illegible]

PostgreSQL[illegible]

[illegible]13[illegible] [illegible]

[illegible]DBA[illegible]DBA[illegible]PostgreSQL[illegible]

13.1 [illegible]

[illegible]DRBD[illegible]

·[illegible]

·[illegible]“[illegible]”[illegible]TRUNCATE[illegible]ALTER TABLE DROP COLUMN[illegible]

·[illegible]

·[illegible]BUG[illegible]WHERE[illegible]UPDATE[illegible]DELETE[illegible]

·[illegible]

[illegible]

[illegible]CREATE TABLE AS[illegible]COPY[illegible]SQL[illegible]

[illegible]24[illegible]

[illegible]DBA[illegible]RTO[illegible]Recovery Time Objective[illegible]RPO[illegible]Recovery Point Objective[illegible]RTO[illegible]RPO[illegible]RTO[illegible]RPO[illegible]

[illegible]RPO[illegible]RTO[illegible]

[illegible]CPU[illegible]

[illegible]CPU[illegible]

PostgreSQL□□□SQL□□□PostgreSQL□□□

·□□pg_dump□pg_dumpall□□□□□□SQL□□□□□□□

·□□□□□□□□□

·□□□□□□□□□□□□□PITR□

□□□□□□□□□□□□□□□□

13.2 □□□□

PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□WAL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□WAL□□□WAL□□□□□□□□□□□□□□□□□□□□□□WAL□□□□□□□□□□□□□□□□□□□□□N□□□□□N=2×checkpoint_segment+1□□□□□PostgreSQL□□□□□□□□□WAL□□□□□□□□□□□□□□□□□256□WAL□□□□□□□□□□□□1□□□2□□□3□□□□WAL□□□□□WAL□□□□□□□□□□□□□□□□□□□□□□□□□□□□1□WAL□□□□□□□□□WAL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□WAL□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□WAL□□□WAL□□□DBA□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□DBA□□□DBA□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□

13.2.1 □□WAL□□

1.□□□□□□

□□data□□□□□□□backups□scripts□archive_wals□□□□□□□□□□□

```
[root@pghost1 ~]$ mkdir -p
/pgdata/10/{data,backups,scripts,archive_wals}
[root@pghost1 ~]$ chown -R postgres.postgres /pgdata/10
```

□□□data□□□□□□□□□□□□□□backups□□□□□□□□□□□□□□scripts□□□□□□□□□□□□□□□archive_wals□□□□□□□□□□□□□□□□□□□□□□□NFS□□□□□□□□□□□□□□□□□postgres□□□□□

2.□□wal_level□□

wal_level□□□□□□□minimal□replica□logical□□minimal□replica□□logical□□□WAL□□□□□□□□□WAL□□□□□□□□□□□□minimal□□□□□WAL□□□□□□□□□□WAL□□□□□□□□□□□□□□

minimal□□□□□□□archive_mode□□□□□WAL□□wal_level□□□□□replica□□□□□□

```
mydb=# ALTER SYSTEM SET wal_level = 'replica';
ALTER SYSTEM
```

3.□□archive_mode□□

archive_mode□□□□□□□on□off□always□□□□□off□□□□□□□□□□on□□□□□□

```
mydb=# ALTER SYSTEM SET archive_mode = 'on';
ALTER SYSTEM
```

□□□□□□□□□□□□□□□□□□□

4.□□archive_command□□

archive_command□□□□□□□□□□□□□□□□□□□□□shell□□□□□□□□□shell□□□□archive_command□shell□□□□□□□□□“%p”□□□□□□□WAL□□□□□□□□□□□□□□□□□“%f”□□□□□□□□□□WAL□□□□□□□

□□□□□□archive_command□□□□□archive_command='cp%p/pgdata/10/archi

ve_wals/%f'[illegible]

[illegible]wal_level[illegible]archive_mode[illegible]archive_command[illegible]reload[illegible]archive_command[illegible]WAL[illegible]WAL[illegible]pg_wal[illegible]pg_wal[illegible]wal_level[illegible]archive_mode[illegible]archive_command[illegible]/bin/true[illegible]archive_command[illegible]reload[illegible]

[illegible]WAL[illegible]gzip[illegible]bzip2[illegible]lz4[illegible]archive_command[illegible]pg_wal[illegible]lz4[illegible]WAL[illegible]/pgdata/10/archive_wals/[illegible]

```
mydb=# ALTER SYSTEM SET archive_command = '/usr/bin/lz4 -q -
z %p /pgdata/10/archive_wals/%f.lz4';
ALTER SYSTEM
mydb=# SELECT pg_reload_conf();
    pg_reload_conf
----------------
```

```
     t
(1 row)
mydb=#  show archive_command ;
        archive_command
------------------------------------------------------
    /usr/bin/lz4 -q -z %p /pgdata/10/archive_wals/%f.lz4
(1 row)
```

13.2.2 □□□□□□

□□□□PostgreSQL□□□□□□pg_start_backup□pg_stop_backup□□□□API□□□□□□□PostgreSQL 9.1□□□□□pg_basebackup□□□□□□□□□□□□□□□□□pg_basebackup□□□□□□□□tar□□□□□□□□□□□□□□□□□□□pg_start_backup□pg_stop_backup□□□□□□□□□□□□□□□□□□□□□□□□□□□□□rsync□scp□□□□□□□□□□□□□□□□□□API□□□□□□□□□□□API□□□□□□□□□□PIRT□□□□□□□□□□□□□□□□□□□□□

1.□□□□API□□□□□□

□□□□API□□□□□□□□□□□□□□□□pg_start_backup□□□□□□□□□□□□□□□□□□□□□□□□□□pg_stop_backup□□□□□□□

□□1 □□pg_start_backup□□□

pg_start_backup[illegible]

1[illegible]WAL[illegible]

[illegible]WAL[illegible]

```
NOTICE: WAL archiving is not enabled; you must ensure that
all required WAL segments are copied through other means to
complete the backup
```

[illegible]WAL[illegible]
WAL[illegible]WAL[illegible]
pg_start_backup[illegible]WAL[illegible]
WAL[illegible]13.2.1[illegible]

2[illegible]

[illegible]full_page_writes[illegible]
off[illegible]full_page_writes[illegible]off[illegible]full_page_writes[illegible]on[illegible]

3[illegible]

4□□□□□□□□□□□□□□□backup_label□□□□□backup_label□□□□□□□□□

·START WAL LOCATION□25/2B002118□file 00000001000000250000002B□

·CHECKPOINT LOCATION□□□□□□□□□□□□□LSN□□□

·BACKUP METHOD□□□□□□□□□□□□pg_start_backup□□pg_basebackup□□□□□□□□□□□BACK up METHOD□□□streamed□

·BACKUP FROM□□□□□□□□□master□standby□□□□□□□

·START TIME□□□pg_start_backup□□□□□

·LABEL□□pg_start_backup□□□□□□□

□□□□□□pg_start_backup□□□□□□

```
pg_start_backup(label text [, fast boolean [, exclusive
boolean ]])
```

□□□□□□□□□□□□□□□□□□□label□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

pg_start_backup□□□□□□□CHECKPOINT□□□fast□□□□□□false□□□□□□□□□□□□exclusive□□□□pg_start_backup□□□false□□□□□□□□□□□□□pg_is_in_backup□□exclusive□□□□□TRUE□□□□□□□□□□□□□□□□□□□□□□□

□□2 □□□□□□□□□□□□□□

□□rsync□tar□cp□scp□□□□□□□□□□□□□□□□□□□□□□□□□□pg_wal□pg_replslot□□□postmaster.opts□□□postmaster.pid□□□□□□□□□□□□□□□□□□□

□□3 □□pg_stop_backup□□□

□□□pg_stop_backup□□□□□□□□□□□□□□□□

·□□□□□pg_start_backup□□□□full-page-writes□□□□□□□□□□□□□□□pg-stop-backup□□□□□□□

·□□□□□□□□XLOG□□□

·□□WAL□□□□

·□□□□□□□□□□□□□□□□backup_label□□□□□□□□□pg_stop_backup□□□□□

·□□backup_label□□□backup_label□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□1　□□pg_start_backup□□□□□□□□□□

```
mydb=# SELECT pg_start_backup('base',false,false);
    pg_start_backup
-----------------
    0/4000028
(1 row)
```

□□2　□□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ cd /pgdata/10/backups
[postgres@pghost1 /pgdata/10/backup]$ tar -cvf base.tar.gz
/pgdata/10/data  --exclude=/pgdata/10/data/postmaster.pid --
exclude=/pgdata/10/data/postmaster.opts --
exclude=/pgdata/10/data/log/*
tar: Removing leading `/' from member names
/pgdata/10/data/
...
...
...
/pgdata/10/data/pg_wal/000000010000000000000002.00000028.bac
kup
[postgres@pghost1 /pgdata/10/backup]$
[postgres@pghost1 /pgdata/10/backup]$ ll -h
```

```
total 80M
-rw-r--r-- 1 postgres postgres 80M Feb 13 00:35 base.tar.gz
```

□□3　□□pg_stop_backup□□□□□□□□□

```
mydb=# SELECT pg_stop_backup(false);
NOTICE:  pg_stop_backup complete, all required WAL segments
have been archived
        pg_stop_backup
------------------------------------------------------------
---------------
    (0/4000168,"START WAL LOCATION: 0/4000028 (file
000000010000000000000004) +
    CHECKPOINT LOCATION: 0/4000098
+
    BACKUP METHOD: streamed
+
    BACKUP FROM: master
+
    START TIME: 2018-02-13 00:34:24 CST
+
    LABEL: base
+
    ","")
(1 row)
```

□□□□□□□□□□□□□□□□□□

2.□□pg_basebackup□□□□□□

pg_basebackup□□□□□□□□□□12□□□□□□□□□□□□□□□□□□□□□□pg_basebackup□□□□□□□□□□

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_basebackup -Ft -
Pv -Xf -z -Z5 -p 1922 -D /pgdata/10/backups/
pg_basebackup: initiating base backup, waiting for
checkpoint to complete
pg_basebackup: checkpoint completed
pg_basebackup: write-ahead log start point: 33/A8000028 on
timeline 1
52782/52782 kB (100%), 1/1 tablespace
pg_basebackup: write-ahead log end point: 33/A8000130
pg_basebackup: base backup completed
```

□□□□□□□□□base backup completed□□□□□□□□□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ ll -h /pgdata/10/backups/
total 3.8M
-rw-r--r-- 1 postgres postgres 3.8M Feb 11 01:26 base.tar.gz
```

□□□□□□□□□□□□□□□□□□

13.3 [illegible]

[illegible]WAL[illegible]

[illegible]

```
CREATE TABLE tbl
(
    id SERIAL PRIMARY KEY,
    ival INT NOT NULL DEFAULT 0,
    description TEXT,
    created_time TIMESTAMPTZ NOT NULL DEFAULT now()
);
```

[illegible]

```
mydb=# INSERT INTO tbl (ival) VALUES (1);
INSERT 0 1
mydb=# SELECT id,ival,description,created_time FROM tbl;
    id | ival | description |          created_time
-------+------+-------------+-------------------------------
     1 |    1 |             | 2018-02-13 01:26:36.767887+08
(1 row)
```

[illegible]WAL[illegible]16MB[illegible]WAL[illegible]

```
mydb=# SELECT pg_switch_wal();
    pg_switch_wal
---------------
    0/3000350
(1 row)
```

□□□□□□archive_timeout□□□□□□timeout□□□□□□□□□□WAL□□

13.3.1 □□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□ival□□2□□□□□□□□□□□

```
mydb=# INSERT INTO tbl (ival) VALUES (2);
INSERT 0 1
mydb=# SELECT id,ival,description,created_time FROM tbl;
    id | ival | description |         created_time
-------+------+-------------+-------------------------------
     1 |    1 |             | 2018-02-13 01:26:36.767887+08
     2 |    2 |             | 2018-02-13 01:47:42.280831+08
(2 rows)
```

□□□□□□□□□created_time□□□□01□47□42.280831+08□

□□□□□□□□□□□□□

1□□

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_ctl -D
/pgdata/10/data/ -mi stop
waiting for server to shut down.... done
server stopped
[postgres@pghost1 ~]$ rm -rf /pgdata/10/data
```

2[illegible]pg_basebackup[illegible]

```
[postgres@pghost1 ~]$ mkdir -p /pgdata/10/data
[postgres@pghost1 ~]$ chmod 0700 /pgdata/10/data
[postgres@pghost1 ~]$ tar -xvf /pgdata/10/backups/base.tar.gz
-C /pgdata/10/data/
```

3[illegible]recovery.conf[illegible]

[illegible]share[illegible]recovery.conf[illegible]recovery.conf[illegible]0600[illegible]

```
[postgres@pghost1 ~]$ cp /usr/pgsql-
10/share/recovery.conf.sample /pgdata/10/data/recovery.conf
[postgres@pghost1 ~]$ chmod 0600 /pgdata/10/data/recovery.conf
[postgres@pghost1 ~]$ ll /pgdata/10/data/recovery.conf
-rw------- 1 postgres postgres 5762 Feb 13 01:07
/pgdata/10/data/recovery.conf
```

[illegible]recovery.conf[illegible]

```
[postgres@pghost1 ~]$ vim /pgdata/10/data/recovery.conf
```

□□□□□□□□□□□□□□□resotre_command□□□□□□□□□□

```
restore_command = '/usr/bin/lz4 -d
/pgdata/10/archive_wals/%f.lz4 %p'
```

□□□□□□□□□□□□□□□□□□□recovery_target_timeline□□□□□□□latest□□□□□□□□□□□□□□□□□□□□□□□recovery.conf□□□□□

```
restore_command = '/usr/bin/lz4 -d
/pgdata/10/archive_wals/%f.lz4 %p'
recovery_target_timeline = 'latest'
```

4□□□□□□□□□□□

□□□□□□□□□□□recovery.conf□□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_ctl -D
/pgdata/10/data start
[postgres@pghost1 ~]$
[postgres@pghost1 ~]$ tailf /pgdata/10/data/log/postgresql-
Tue.csv
...
...
...
2018-02-13 01:32:06.745 CST,,,41781,,5a81cf96.a335,2,,2018-02-
13 01:32:06 CST,,0,LOG,00000,"starting archive
recovery",,,,,,,,,""
...
...
```

```
...
2018-02-13 01:32:06.908 CST,,,41781,,5a81cf96.a335,9,,2018-02-
13 01:32:06 CST,1/0,0,LOG,00000,"redo done at
0/3000060",,,,,,,,,""
2018-02-13 01:32:06.946 CST,,,41781,,5a81cf96.a335,10,,2018-
02-13 01:32:06 CST,1/0,0,LOG,00000,"restored log file
""000000010000000000000003"" from archive",,,,,,,,,""
2018-02-13 01:32:06.961 CST,,,41781,,5a81cf96.a335,11,,2018-
02-13 01:32:06 CST,1/0,0,LOG,00000,"selected new timeline ID:
2",,,,,,,,,""
2018-02-13 01:32:07.007 CST,,,41781,,5a81cf96.a335,12,,2018-
02-13 01:32:06 CST,1/0,0,LOG,00000,"archive recovery
complete",,,,,,,,,""
2018-02-13 01:32:07.413 CST,,,41779,,5a81cf95.a333,3,,2018-02-
13 01:32:05 CST,,0,LOG,00000,"database system is ready to
accept connections",,,,,,,,,""
...
...
...
```

□□□□□□□□recovery.conf□□PostgreSQL□□□□□□recovery.done□□□□□□□□□□□□□

5□□□□□□□□□□□□□□□□□□□

```
mydb=# SELECT id,ival,description,created_time FROM tbl;
    id | ival | description |         created_time
-------+------+-------------+-------------------------------
     1 |    1 |             | 2018-02-13 01:26:36.767887+08
     2 |    2 |             | 2018-02-13 01:47:42.280831+08
(2 rows)
```

□□□□□□□□□□□□□□□□□□□□□□□□□

13.3.2 □□□□□□□

[illegible]

[illegible]

```
mydb=# SELECT id,ival,description,created_time FROM tbl ORDER
BY created_time DESC;
    id | ival | description |         created_time
-------+------+-------------+-------------------------------
     3 |    3 |             | 2018-02-28 12:20:05.003997+08
     2 |    2 |             | 2018-02-28 12:14:05.248928+08
     1 |    1 |             | 2018-02-28 12:13:59.472933+08
(3 rows)
mydb=# SELECT current_timestamp;
       current_timestamp
-------------------------------
    2018-02-28 15:12:12.185701+08
(1 row)
```

[illegible]WHERE[illegible]DELETE[illegible]

```
mydb=# DELETE FROM tbl;
DELETE 3
mydb=# SELECT id,ival,description,created_time FROM tbl ORDER
BY created_time DESC;
    id | ival | description | created_time
-------+------+-------------+--------------
(0 rows)
```

[illegible]current_timestamp[illegible]recovery.conf[illegible]

```
restore_command = '/usr/bin/lz4 -d
/pgdata/10/archive_wals/%f.lz4 %p'
recovery_target_time = '2018-02-28 15:12:12.185701+08'
```

```
2018-02-28 15:19:42.025 CST,,,16251,,5a96580e.3f7b,2,,2018-02-
28 15:19:42CST,, 0,LOG,00000, "starting point-in-time recovery
to 2018-02-28 15:12:12.185701+ 08",,,,,,,,,""
2018-02-28 15:19:42.203 CST,,,16251,,5a96580e.3f7b,9,,2018-02-
28 15:19:42 CST,1/0,0,LOG,00000,"selected new timeline ID:
4",,,,,,,,,""
...
...
...
2018-02-28 15:19:42.678 CST,,,16249,,5a96580d.3f79,3,,2018-02-
28 15:19:41 CST,,0,LOG,00000,"database system is ready to
accept connections",,,,,,,,,""
```

```
mydb=# SELECT id,ival,description,created_time FROM tbl ORDER
BY created_time DESC;
    id | ival | description |         created_time
-------+------+-------------+-------------------------------
     3 |    3 |             | 2018-02-28 12:20:05.003997+08
     2 |    2 |             | 2018-02-28 12:14:05.248928+08
     1 |    1 |             | 2018-02-28 12:13:59.472933+08
(3 rows)
```

13.3.3

[illegible]

[illegible]pg_create_restore_point[illegible]

```
mydb=# \df pg_create_restore_point
                                   List of functions
    Schema    |       Name             | Result data type |
Argument data types| Type
--------------+---------------------+---------------+-------
------------------
    pg_catalog | pg_create_restore_point | pg_lsn          |
text             | normal
(1 row)
```

[illegible]

```
mydb=# INSERT INTO tbl (ival) VALUES (4);
INSERT 0 1
mydb=# SELECT id,ival,description,created_time FROM tbl ORDER
BY created_time DESC;
    id | ival | description |         created_time
-------+------+-------------+-------------------------------
     4 |    4 |             | 2018-02-28 15:21:21.960729+08
     3 |    3 |             | 2018-02-28 12:20:05.003997+08
     2 |    2 |             | 2018-02-28 12:14:05.248928+08
     1 |    1 |             | 2018-02-28 12:13:59.472933+08
(4 rows)
```

[illegible]

```
mydb=# SELECT pg_create_restore_point('restore_point');
    pg_create_restore_point
-------------------------
    0/C0000C8
(1 row)
```

□□□□□□□□□□□□□□DROP□□□□□□□□□□

```
mydb=# DELETE FROM tbl WHERE id = 4;
DELETE 1
mydb=# ALTER TABLE tbl DROP COLUMN description;
ALTER TABLE
mydb=# SELECT * FROM tbl ORDER BY created_time DESC;
    id | ival |         created_time
-------+------+-------------------------------
     3 |    3 | 2018-02-28 12:20:05.003997+08
     2 |    2 | 2018-02-28 12:14:05.248928+08
     1 |    1 | 2018-02-28 12:13:59.472933+08
(3 rows)
```

□□□□□□□□□“restore_point”□□□□□□□□□□□

```
[postgres@pghost1 /pgdata/10/data]$ /usr/pgsql-10/bin/pg_ctl -
D /pgdata/10/data -mi stop
waiting for server to shut down.... done
server stopped
[postgres@pghost1 /pgdata/10/data]$
[postgres@pghost1 /pgdata/10/data]$ rm -rf *
[postgres@pghost1 /pgdata/10/data]$ tar -xf
/pgdata/10/backups/base.tar.gz -C /pgdata/10/data/
[postgres@pghost1 /pgdata/10/data]$ cp /usr/pgsql-
10/share/recovery.conf.sample /pgdata/10/data/recovery.conf
[postgres@pghost1 /pgdata/10/data]$ chmod 0600
/pgdata/10/data/recovery.conf
[postgres@pghost1 /pgdata/10/data]$ ll
/pgdata/10/data/recovery.conf
```

```
-rw------- 1 postgres postgres 5762 Feb 13 01:07
/pgdata/10/data/recovery.conf
```

□□recovery.conf□□□□□□

```
[postgres@pghost1 ~]$ vim /pgdata/10/data/recovery.conf
```

□□□□□□□□□□□□□□restore_command□recovery_target_name□□□□□□□□

```
restore_command = '/usr/bin/lz4 -d
/pgdata/10/archive_wals/%f.lz4 %p'
recovery_target_name = 'restore_point'
```

□□□□□□□□□□□□□□□□□□□□□□□□

```
...
...
...
2018-02-28 16:28:33.396 CST,,,32457,,5a966831.7ec9,2,,2018-02-
28 16:28:33 CST,,0,LOG,00000,"starting point-in-time recovery
to ""restore_point""",,,,,,,,,""
...
...
...
2018-02-28 16:28:33.450 CST,,,32455,,5a966830.7ec7,2,,2018-02-
28 16:28:32 CST,,0,LOG,00000,"database system is ready to
accept read only connections",,,,,,,,,""
...
...
...
2018-02-28 16:28:33.985 CST,,,32455,,5a966830.7ec7,3,,2018-02-
28 16:28:32 CST,,0,LOG,00000,"database system is ready to
accept connections",,,,,,,,,""
```

```
...
...
...
```

[illegible]

```
mydb=# SELECT * FROM tbl ORDER BY created_time DESC;
    id | ival | description |         created_time
-------+------+-------------+-------------------------------
     4 |    4 |             | 2018-02-28 15:21:21.960729+08
     3 |    3 |             | 2018-02-28 12:20:05.003997+08
     2 |    2 |             | 2018-02-28 12:14:05.248928+08
     1 |    1 |             | 2018-02-28 12:13:59.472933+08
(4 rows)
```

[illegible]restore_point[illegible]

13.3.4 [illegible]

PostgreSQL[illegible]

[illegible]

```
mydb=# SELECT * FROM tbl ORDER BY created_time DESC;
    id | ival | description |         created_time
-------+------+-------------+-------------------------------
     4 |    4 |             | 2018-02-28 15:21:21.960729+08
     3 |    3 |             | 2018-02-28 12:20:05.003997+08
     2 |    2 |             | 2018-02-28 12:14:05.248928+08
     1 |    1 |             | 2018-02-28 12:13:59.472933+08
(4 rows)
```

[illegible]

```
mydb=# BEGIN;
BEGIN
mydb=# SELECT txid_current();
    txid_current
--------------
        561
(1 row)
mydb=# DELETE FROM tbl WHERE id > 1;
DELETE 3
mydb=# END;
COMMIT
```

[illegible]

```
mydb=# SELECT * FROM tbl ORDER BY created_time DESC;
    id | ival | description |         created_time
-------+------+-------------+-------------------------------
     1 |    1 |             | 2018-02-28 12:13:59.472933+08
(1 row)
```

[illegible]xid[illegible]561[illegible]561[illegible]recovery.conf[illegible]recovery.conf[illegible]

```
restore_command = '/usr/bin/lz4 -d
/pgdata/10/archive_wals/%f.lz4 %p'
recovery_target_xid = 561
```

[illegible]restore_command[illegible]recovery_target_xid[illegible]

[illegible]

```
...
...
...
2018-02-28 18:06:55.830 CST,,,7036,,5a967f3f.1b7c,2,,2018-02-
28 18:06:55 CST,,0,LOG,00000,"starting point-in-time recovery
to XID 561",,,,,,,,,""
    2018-02-28 18:06:56.003 CST,,,7036,,5a967f3f.1b7c,9,,2018-
02-28 18:06:55 CST,1/0,0,LOG,00000,"selected new timeline ID:
7",,,,,,,,,""
...
...
...
2018-02-28 18:06:55.883 CST,,,7033,,5a967f3e.1b79,2,,2018-02-
28 18:06:54 CST,,0,LOG,00000,"database system is ready to
accept read only connections",,,,,,,,,""
...
...
...
```

[illegible]

```
mydb=# SELECT * FROM tbl ORDER BY created_time DESC;
    id | ival | description |         created_time
-------+------+-------------+------------------------------
     4 |    4 |             | 2018-02-28 15:21:21.960729+08
     3 |    3 |             | 2018-02-28 12:20:05.003997+08
     2 |    2 |             | 2018-02-28 12:14:05.248928+08
     1 |    1 |             | 2018-02-28 12:13:59.472933+08
(4 rows)
```

[illegible]

13.3.5 [illegible]

□□□13.3.1□□□□□□□□□□□□□recovery_target_timeline□□□□□□□□timeline□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□initdb□□□□□□□□□□TimeLineID□1□□□□□□□□□□□TimeLineID□□□□□□

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/initdb -D
/pgdata/10/data
...
...
...
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_controldata
/pgdata/10/data | grep TimeLineID
Latest checkpoint's TimeLineID:       1
Latest checkpoint's PrevTimeLineID:   1
```

□□□□□□□□TimeLineID□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□TimeLineID□□□□□□

```
mydb=# INSERT INTO tbl (ival,description) VALUES (1,'□□□□□□
□');
INSERT 0 1
mydb=# INSERT INTO tbl (ival,description) VALUES (2,'□□□□□□');
INSERT 0 1
mydb=# SELECT id,ival,description,created_time FROM tbl;
    id | ival |  description  |         created_time
-------+------+---------------+------------------------------
     1 |    1 | □□□□□□□  | 2018-02-28 21:23:41.881687+08
     2 |    2 | □□□□□□    | 2018-02-28 21:24:25.07764+08
(2 rows)
mydb=# SELECT pg_switch_wal();
    pg_switch_wal
```

```
---------------
    0/167D4D8
(1 row)
```

[illegible]pg_wal[illegible]

```
[postgres@pghost1 ~]$ ll -h /pgdata/10/data/pg_wal/
total 33M
-rw------- 1 postgres postgres  16M Feb 28 21:29
000000010000000000000001
-rw------- 1 postgres postgres  16M Feb 28 21:29
000000010000000000000002
drwx------ 2 postgres postgres 4.0K Feb 28 21:29
archive_status
[postgres@pghost1 ~]$ ll -h
/pgdata/10/data/pg_wal/archive_status/
total 0
-rw------- 1 postgres postgres 0 Feb 28 21:29
000000010000000000000001.done
[postgres@pghost1 ~]$ ll -h /pgdata/10/archive_wals/
total 2.3M
-rw------- 1 postgres postgres 2.3M Feb 28 21:29
000000010000000000000001.lz4
```

[illegible]pg_switch_wal[illegible]WAL[illegible]pg_wal[illegible]archive_status[illegible]

[illegible]TimeLineID[illegible]

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_controldata
/pgdata/10/data | grep TimeLineID
Latest checkpoint's TimeLineID:       2
Latest checkpoint's PrevTimeLineID:   1
```

□□□□pg_wal□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ ll -h /pgdata/10/data/pg_wal/
total 33M
-rw------- 1 postgres postgres  16M Feb 28 22:19
000000010000000000000003
-rw------- 1 postgres postgres  16M Feb 28 22:21
000000020000000000000004
-rw------- 1 postgres postgres   41 Feb 28 22:19
00000002.history
drwx------ 2 postgres postgres 4.0K Feb 28 22:19
archive_status
[postgres@pghost1 ~]$ ll -h
/pgdata/10/data/pg_wal/archive_status/
total 0
-rw------- 1 postgres postgres 0 Feb 28 21:33
000000010000000000000003.done
-rw------- 1 postgres postgres 0 Feb 28 22:19
00000002.history.done
```

□□□□pg_wal□□□□□□WAL□□□□□□□□□□□“.history”□□□TimeLineID□TimeLineID□□□□1□□□4□□□□□□□□□□□□□□□□□□□□“TimeLineID.history”□□□□TimeLineID□2□□□□□□□□□□00000002.history□

□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ ll -h /pgdata/10/archive_wals/
total 2.5M
-rw------- 1 postgres postgres 2.3M Feb 28 21:29
```

```
000000010000000000000001.lz4
-rw------- 1 postgres postgres  78K Feb 28 21:33
000000010000000000000002.lz4
-rw------- 1 postgres postgres  216 Feb 28 21:33
000000010000000000000003.00000028.backup.lz4
-rw------- 1 postgres postgres  78K Feb 28 21:33
000000010000000000000003.lz4
-rw------- 1 postgres postgres   60 Feb 28 22:19
00000002.history.lz4
```

□□□□□□□□□□□□□□□□□□

```
mydb=# INSERT INTO tbl (ival,description) VALUES (3,'□□□□□□
□');
INSERT 0 1
mydb=# INSERT INTO tbl (ival,description) VALUES (4,'□□□□□');
INSERT 0 1
mydb=# SELECT id,ival,description,created_time FROM tbl;
    id | ival |  description  |         created_time
-------+------+---------------+-------------------------------
     1 |    1 | □□□□□□□  | 2018-02-28 21:23:41.881687+08
     2 |    2 | □□□□□□    | 2018-02-28 21:24:25.07764+08
     3 |    3 | □□□□□□□  | 2018-02-28 22:21:03.68966+08
     4 |    4 | □□□□□     | 2018-02-28 23:01:58.253911+08
(4 rows)
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□“4|4|□□□□|2018-02-2823□01□58.253911+08”□□□□□□□□□□□□recovery_target_timeline□□2□□□□□□□□□recovery.conf□□□□

```
restore_command = '/usr/bin/lz4 -d
/pgdata/10/archive_wals/%f.lz4 %p'
recovery_target_timeline = 2
recovery_target_time = '2018-02-28 23:00:00'
```

□□□□□□□□□□□□□□□

```
mydb=# SELECT id,ival,description,created_time FROM tbl;
    id | ival |  description   |         created_time
-------+------+----------------+------------------------------
-
     1 |    1 | □□□□□□□   | 2018-02-28 21:23:41.881687+08
     2 |    2 | □□□□□□     | 2018-02-28 21:24:25.07764+08
     3 |    3 | □□□□□□□   | 2018-02-28 22:21:03.68966+08
(3 rows)
```

□□□□□□□□□□□□□2□□□□□□□

13.4 SQL[illegible]

[illegible]

[illegible]PostgreSQL[illegible]SQL[illegible]

13.4.1 SQL[illegible]

[illegible]SQL[illegible]SQL[illegible]PostgreSQL[illegible]pg_dump[illegible]pg_dumpall[illegible]SQL[illegible]pg_dump[illegible]pg_dumpall[illegible]pg_dump[illegible]pg_dumpall[illegible]

1.pg_dump[illegible]pg_dumpall[illegible]

pg_dump[illegible]

```
pg_dump dumps a database as a text file or to other formats.
Usage:
```

```
pg_dump [OPTION]... [DBNAME]
```

pg-dump□□□OPTION□□□□□

```
-F, --format=c|d|t|p
```

□□□□□□dump□□□□□□□□□□□□□□

·c□□□□□□□□□□

·d□□□□□□□□□□□□□□□□□□□□□□

·t□□□tar□

·p□□□□□□SQL□□

```
-j, --jobs=NUM
```

□□□□□□□□□□F□□□□□d□□□□□□□□□□□□□□□dump□

```
-a, --data-only
```

□dump□□□□□□

```
-c, --clean
```

□□□□□DROP□□□□□□□□

```
-C, --create
```

□□□□Database□□□□

```
-n, --schema=SCHEMA
-N, --exclude-schema=SCHEMA
```

□□Schema□□□□□□Schema□

```
-s, --schema-only
-t, --table=TABLE
-T, --exclude-table=TABLE
```

□□□□□□□□□□□□□□□Schema□□□

```
--inserts                    dump data as INSERT commands,
rather than COPY
--column-inserts             dump data as INSERT commands
with column names
```

□□□□pg_dump□□□SQL□□□□□PostgreSQL□□□COPY□□□□□□□□□□□□□□□□□□SQL□□□□

INSERT[illegible]INSERT[illegible]PostgreSQL[illegible]--insert[illegible]--column-inserts[illegible]INSERT[illegible]

[illegible]SQL[illegible]INSERT[illegible]

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_dump -Fp -a --
insert --column-inserts -t tbl -p 1921 mydb > dump.sql
[postgres@pghost1 ~]$ cat dump.sql
--
-- PostgreSQL database dump
--
-- Dumped from database version 10.2
-- Dumped by pg_dump version 10.2
SET statement_timeout = 0;
...
...
...
SET search_path = public, pg_catalog;
--
-- Data for Name: tbl; Type: TABLE DATA; Schema: public;
Owner: postgres
--
INSERT INTO tbl (id, ival, description, created_time) VALUES
(1, 1, '[illegible]', '2018-02-28 21:23:41.881687+08');
INSERT INTO tbl (id, ival, description, created_time) VALUES
(2, 2, '[illegible]', '2018-02-28 21:24:25.07764+08');
INSERT INTO tbl (id, ival, description, created_time) VALUES
(3, 3, '[illegible]', '2018-02-28 22:21:03.68966+08');
--
-- Name: tbl_id_seq; Type: SEQUENCE SET; Schema: public;
Owner: postgres
--
SELECT pg_catalog.setval('tbl_id_seq', 36, true);
--
```

```
-- PostgreSQL database dump complete
--
```

pg_dumpall[illegible]pg_dump[illegible]

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_dumpall -r -p
1921
--
-- PostgreSQL database cluster dump
--
SET default_transaction_read_only = off;
SET client_encoding = 'UTF8';
SET standard_conforming_strings = on;
--
-- Roles
--
CREATE ROLE postgres;
ALTER ROLE postgres WITH SUPERUSER INHERIT CREATEROLE
CREATEDB LOGIN REPLICATION BYPASSRLS;
--
-- PostgreSQL database cluster dump complete
--
```

2.SQL[illegible]

[illegible]psql[illegible]SQL[illegible]

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/psql -p 1922 mydb <
dump.sql
```

□□□□□□□□□□□□□□□□□□□□pg_restore□□□□
□□□□□□

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_dump -Fc -p 1922
mydb > custom.dat
```

□□□□□pg_restore□□□□□□□□

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_restore -p 1922 -
d mydb custom.dat
```

pg_dump□pg_dump□pg_restore□□□□□□
□□□□□□□□□□□□□□□□□□□□□□

13.4.2 □□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□cp□tar□nc□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□

13.5 [illegible]

[illegible] pg_rman [illegible] pg_backrest [illegible]

[illegible]14[illegible] [illegible]

[illegible]DBA[illegible]Oracle[illegible]Oracle[illegible]RAC[illegible]RAC[illegible]PostgreSQL[illegible]RAC[illegible]PostgreSQL[illegible]12[illegible]

PostgreSQL[illegible]WAL[illegible]VIP[illegible]virtual ip[illegible]IP[illegible]VIP[illegible]VIP[illegible]

[illegible]WAL[illegible]

□□□

□□□□□□PostgreSQL□□□□□□□□□□□Pgpool-II+□□□□□□□□□□□□□Keepalived+□□□□□□□□

14.1 Pgpool-II+[illegible]

[illegible]Pgpool-II[illegible]Pgpool-II[illegible]Pgpool-II[illegible]Pgpool-II[illegible]pgpool[illegible]

·[illegible]pgpool[illegible]

·[illegible]SELECT[illegible]pgpool[illegible]

·[illegible]pgpool[illegible]

·[illegible]pgpool[illegible]PostgreSQL[illegible]pgpool[illegible]PostgreSQL[illegible]

[illegible]pgpool[illegible]pgpool[illegible]pgpool[illegible]pgpool[illegible]

□□□□□□pgpool□□□□□□□□□□□pgpool□□□□
□□□□□□□□□□□□□

·□□□□□□□□PostgreSQL□□□□□□
PostgreSQL□□□□□pgpool□□□□□□□□□□□□□□□□
□□$prefix/etc/pgpool.conf.sample-stream□□
□□□□□□□□□□

·□□□□□□□□□□□□Slony□pgpool□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□Slony□□□□□□
$prefix/etc/pgpool.conf.sample-master-
slave□□□□□□□□□□□□

·□□□□□□□□□□□□pgpool□□□□□□□□□□□□
pgpool□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□
$prefix/etc/pgpool.conf.sample-
replication.□□□□□□□□□□□□

·□□□□□□□□□pgpool□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□
$prefix/etc/pgpool.conf.sample□□□□□□□□□□
□□□

Pgpool□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□

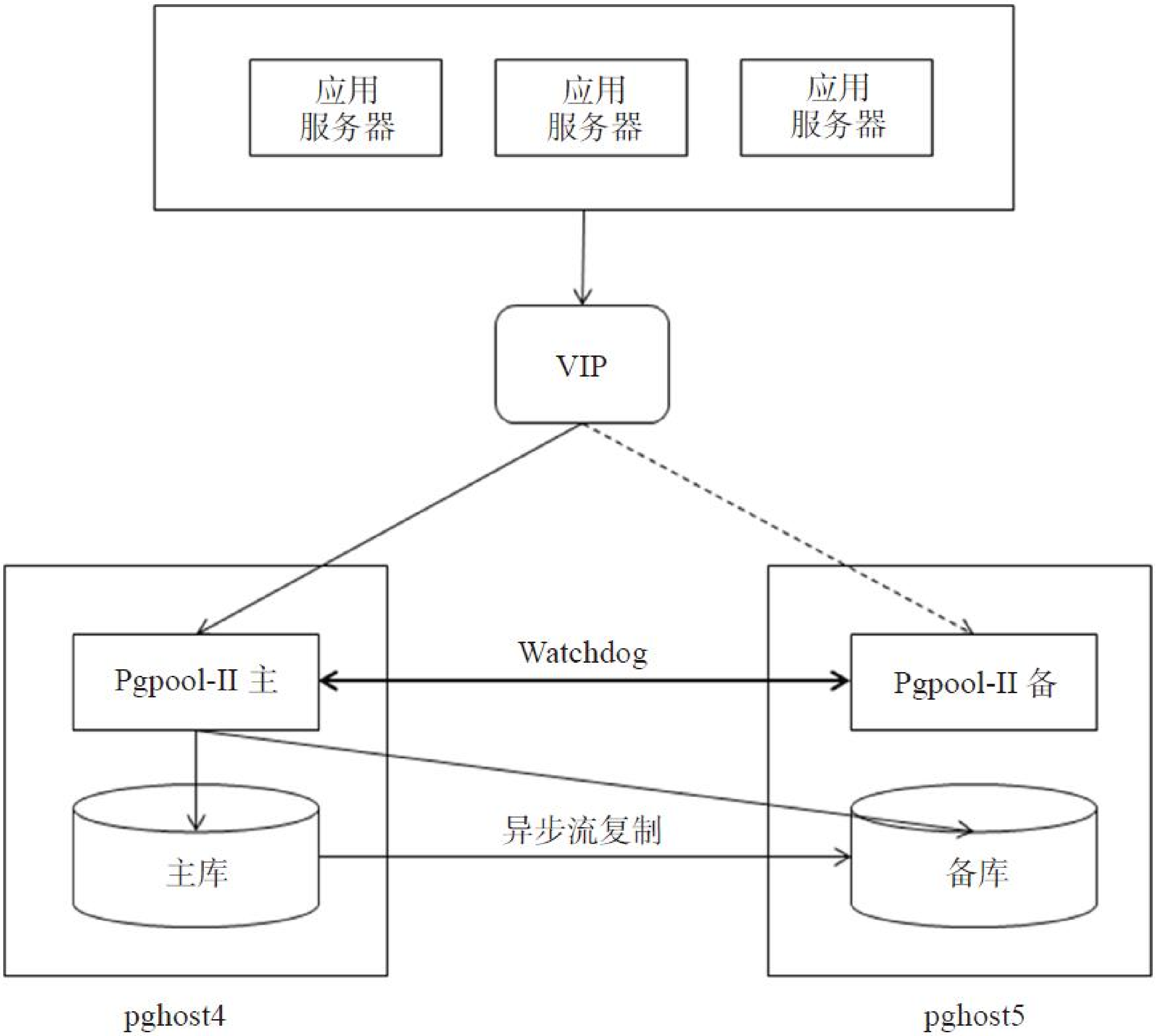
□□□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□12.1□□□□□□□□14-1□

表14-1　pgpool+□□□□□□□□□

主机名	组件	IP地址	端口	版本	VIP
pghost4	主库	192.168.26.57	1921	PostgreSQL10	192.168.26.72
	pgpool主	192.168.26.57	9999	Pgpool-II 3.6.6	
pghost5	备库	192.168.26.58	1921	PostgreSQL10	
	pgpool备	192.168.26.58	9999	Pgpool-II 3.6.6	

14.1.1　pgpool□□□□□

pgpool□□□□14-1□□□



□14-1 pgpool□□□□□

pghost4□pghost5□□□□□□□□□□pghost4□□□□□□□□pghost5□□□□□□□□□□pgpool□□□□□□pghost4□□pgpool□□□□□□pghost5□□□□□□□□□□watchdog□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

watchdog□pgpool□□□□□□watchdog□pgpool□□□□□□□□□□□□□□□□□pgpool□□□□watchdog□□□□□□□□□□

·□pgpool□□PostgreSQL□□□□□□□□□□pgpool□□□□□□□

·□□□pgpool□□□□□□□□□□□□

·□watchdog□□□□□□□□□□□pgpool□□□□□□pgpool□□□□□□□□□□□□□□□□□□□□pgpool□□□□□□□

·□pgpool□□□□□□□watchdog□□□□□□□□watchdog□□□□

·□pgpool□□□□□□□□□□□□watchdog□□□VIP□□□□□pgpool□□□

14.1.2 pgpool□□

pgpool□□□□□□□□□Pgpool-II 3.6.6□□□□□RPM□□□□□□□□□□□□□□□□□□□□□□□□□□□□□http://www.pgpool.net/download.php?f=pgpool-II-3.6.6.tar.gz □

pghost4□pghost5□□□□pgpool-II-3.6.6.tar.gz□□□□□

```
# tar xvf pgpool-II-3.6.6.tar.gz
```

□□□□□□

```
# ./configure --prefix=/opt/pgpool --with-pgsql=/opt/pgsql
# make
# make install
```

□□□□□/opt/pgpool□□□□□□□□□□□□□□

```
[root@pghost4 etc]# ll /opt/pgpool/
total 20
drwxr-xr-x 2 root root 4096 Oct 4 14:39 bin
drwxr-xr-x 2 root root 4096 Oct 4 19:13 etc
drwxr-xr-x 2 root root 4096 Oct 4 14:39 include
drwxr-xr-x 2 root root 4096 Oct 4 14:39 lib
drwxr-xr-x 3 root root 4096 Oct 4 14:39 share
```

□□/opt/pgpool/bin□□□□pgpool□□□□□□□□pgpool□pg_md5□pcp_attach_node□□/opt/pgpool/etc□□□□□□□□□□□pcp.conf.sample□pgpool.conf.sample-stream□□□□□□□

□□□root□□□□□□□□pgpool□□□□□□□□□□□□□□/opt/pgpool□pgpool□□□□□□□□□□□□□□□□□

1.pghost4□pghost5□□□□/etc/hosts

□□□/etc/hosts□□□□□□□□□□

```
192.168.26.57 pghost4
192.168.26.58 pghost5
```

2.□□pghost4□pghost5□□□□□□□□

□□□root□□□□pgpool□□□□□□□□□□□□□□□□□□pgpool□□□pgpool□□□□□ failover_command□□□□□□□□□□□□ssh□□□□□□□□□□□□□□□

□□□pghost4□pghost5□□□postgres□□□□□□□□□□□□pghost4□□□□□□□□

```
# su - postgres
$ ssh-keygen
$ ssh-copy-id postgres@pghost5
```

□□□pghost4□□□□□□postgres□□□□□□□□□□□□□ssh□pghost5□□□□□□

```
ssh postgres@pghost5
```

□□□□□□□□□pghost5□□□□□□□□□□□

□□□□□□□□□□□□□root□□pgpool□

3.pghost4□pghost5□□□□pool_hba.conf□□□□

□□□□PostgreSQL□□□ $PGDATA/pg_hba.conf□□□□□□□□□□□□□□ pgpool□□□□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□pgpool□pgpool□□□□□□□□□□□□□□□ pgpool□□□□□□□□□□□□□□pgpool□□□□□□□

□□□□□□□□pool_hba.conf.sample□□□□□□□□□pool_hba.conf□□□□□□

```
# cd /opt/pgpool/etc
# cp pool_hba.conf.sample pool_hba.conf
```

□□□pool_hba.conf□□□□□□□□□□□□□□ PostgreSQL□□□□pg_hba.conf□□□ pool_hba.conf□□□□□□□□□□□□□□□

```
host    replication     repuser      192.168.26.57/32
md5
host    replication     repuser      192.168.26.58/32
md5
```

```
host    replication      repuser       192.168.26.72/32
md5
host    all      all      0.0.0.0/0        md5
```

4.pghost4□pghost5□□□□pool_passwd□□□□

□□pgpool□□□MD5□□□□□□□pgpool□□□□pool_passwd□□□□□□□pool_passwd□□□□□□□□□

```
username: encrypted_passwd
```

□□□□□□□□□□□□pgpool□□pg_md5□□□□□□□□□□□□□□

```
# pg_md5 -u postgres -m postgres123
```

pg_md5□□□□□□□□□□□MD5□□□□□MD5□□□□□pool_passwd□□□□□□□□□

```
# cat pool_passwd
postgres:md5163311300b0732b814a34aabfdfffe62
```

□□□□□□□□□□□□□□□□□□□postgres□□□□□□□□□□□□□□□□□□pool_passwd□□□□□□□□

```
postgres=# SELECT rolpassword FROM pg_authid WHERE
rolname='postgres';
           rolpassword
-------------------------------------
 md5163311300b0732b814a34aabfdfffe62
(1 row)
```

□□□postgres□□□□□□MD5□□□□pool_passwd□□□□□□□□□□□

5.pghost4□pghost5□□□□pgpool.conf□□□□

□□□□□/opt/pgpool/etc□□□□□□□□□□□

```
# cd /opt/pgpool/etc
# cp pgpool.conf.sample-stream  pgpool.conf
```

pgpool.conf□□□□□□□□□□□□□□□□□□□□□

pgpool□□□□□□□□□

```
listen_addresses = '*'
port = 9999
```

·listen_addresses□pgpool□□□□□*□□□□□□□□□

·port=9999□pgpool□□□□□□□□□9999□

pgpool□□PostgreSQL□□□□□□□□□□□□pghost4□pghost5□□□□□□□□□□□□□0□□□□□□□1□

```
backend_hostname0 = 'pghost4'
backend_port0 = 1921
backend_data_directory0 = '/data1/pg10/pg_root'
backend_flag0 = 'ALLOW_TO_FAILOVER'
backend_hostname1 = 'pghost5'
backend_port1 = 1921
backend_data_directory1 = '/data1/pg10/pg_root'
backend_flag1 = 'ALLOW_TO_FAILOVER'
```

·backend_hostname0□□□□PostgreSQL□□0□□□□IP□

·backend_port0□□□PostgreSQL□□0□□□□

·backend_data_directory0□□□PostgreSQL□□0□□□□□

·backend_flag0□□□□□□□□□□□□□□□ALLOW_TO_FAILOVER□□□□□□□□□

·backend_hostname1□□□□PostgreSQL□□1□□□□IP□

·backend_port1□□□PostgreSQL□□1□□□□

·backend_data_directory1□□□PostgreSQL□□1□□□□□□

·backend_flag1□□□□□□□□□□□□□□□□□ALLOW_TO_FAILOVER□□□□□□□□□□

pgpool□□□□□□□□□

```
enable_pool_hba = on
pool_passwd = 'pool_passwd'
```

·enable_pool_hba□□□pgpool□□pool_hba.conf□

·pool_passwd□□□MD5□□□□□□□□□□□pool_passwd□

pgpool□□□□□pid□□□□□□□□□

```
log_destination = 'syslog'
pid_file_name = '/opt/pgpool/pgpool.pid'
```

·log_destination□pgpool□□□□□□□□□□□stderr□syslog□□□□□□syslog□/var/log/message□□□□□□□□pgpool□□□

·pid_file_name□pgpool□□□PID□□□

pgpool□□□□□□□□□□□

```
load_balance_mode = off
```

load_balance_mode□□□□□pgpool□□□□□□□□□□□□□SELECT□□□□pgpool□□□□□□□□□□□□□□

pgpool□□□□□□□□□□□□□□□□□□□

```
# pgpool □□□□□□□□□□□
master_slave_mode = on
master_slave_sub_mode = 'stream'
sr_check_period = 10
sr_check_user = 'repuser'
sr_check_password = 're12a345'
sr_check_database = 'postgres'
delay_threshold = 10000000
```

·master_slave_mode□□□□□□□□□□□□off□□□□on□

·master_slave_sub_mode□□□□□□□□□□□slony□stream□slony□□□□slony□□□□□stream□□□□PostgreSQL□□□□□□□□□□□□□stream□

·sr_check_period[illegible]10
[illegible]

·sr_check_user[illegible]

·sr_check_database[illegible]

·delay_threshold[illegible]WAL[illegible]pgpool[illegible]SELECT[illegible]

pgpool[illegible]PostgreSQL[illegible]

```
health_check_period = 5
health_check_timeout = 20
health_check_user = 'repuser'
health_check_password = 're12a345'
health_check_database = 'postgres'
health_check_max_retries = 3
health_check_retry_delay = 3
```

[illegible]

```
failover_command = '/opt/pgpool/failover_stream.sh %d %P %H
%R'
```

failover_command[illegible]pgpool[illegible]pgpool[illegible]%d[illegible]ID[illegible]%P[illegible]ID[illegible]%H[illegible]%R[illegible]failover_stream.sh[illegible]

watchdog[illegible]

```
use_watchdog = on
wd_hostname = 'pghost4'
wd_port = 9000
wd_priority = 1
```

·use_watchdog[illegible]watchdog[illegible]off[illegible]

·wd_hostname[illegible]watchdog[illegible]IP[illegible]pgpool[illegible]

·wd_port[illegible]watchdog[illegible]9000[illegible]

·wd_priority[illegible]watchdog[illegible]pgpool[illegible]watchdog[illegible]pgpool[illegible]pgpool[illegible]

VIP[illegible]

```
delegate_IP = '192.168.26.72'
if_cmd_path = '/sbin'
```

```
if_up_cmd = 'ip addr add $_IP_$/24 dev bond0 label bond0:1'
if_down_cmd = 'ip addr del $_IP_$/24 dev bond0'
```

·delegate_IP□□□pgpool□VIP□pgpool□□□□□□□□VIP□□pgpool□□□□□□□□□□□□pgpool□□□□□□pgpool□□□□□□VIP□

·if_cmd_path□□□□□□□□VIP□□□□□□

·if_up_cmd□□□□□VIP□□□□□□ip addr add□□□□□□VIP□□□pghost4□pghost5□□□□□□□□□□bond0□□□□VIP□□□□□□□□□□bond0□1□□□□□□□□□□□□□□IP□

·if_down_cmd□□□□□VIP□□□□□□ip addr del□□□

watchdog□□□□□□□□□□□

```
heartbeat_destination0 = 'pghost5'    # □□□□pgpool□□□□□□IP
heartbeat_destination_port0 = 9694   # □□□□pgpool□□□□□
heartbeat_device0 = 'bond0'
```

·heartbeat_destination0□□□□□pgpool□□□□□□IP□□□□watchdog□□□□□□pgpool□□□heartbeat_destination□□□□□0□□□

·heartbeat_destination_port0□□□□□pgpool□□□□□□□□□□9694□

·heartbeat_device0□□□pgpool□□watchdog□□□□□□□□□□□□□□□□bond0□

watchdog□□□□□□□□□□□□□□

```
wd_life_point = 3
wd_lifecheck_query = 'SELECT 1'
wd_lifecheck_dbname = 'postgres'
wd_lifecheck_user = 'repuser'
wd_lifecheck_password = 're12a345'
```

·wd_life_point□□□□pgpool□□□□□□□□□□□□

·wd_lifecheck_query□□□pgpool□□□□□SQL□

·wd_lifecheck_dbname□□□pgpool□□□□□□□□□

·wd_lifecheck_user□□□pgpool□□□□□□□□□□□□

·wd_lifecheck_password□□□pgpool□□□□□□□□□□□□□

□□pgpool□□□□□□□□□□□□

```
other_pgpool_hostname0 = 'pghost5'   # □□□□pgpool□□□□□□IP
other_pgpool_port0 = 9999            # □□□□pgpool□□□□□
other_wd_port0 = 9000                # □□□□pgpool□□watchdog□
□□
```

·other_pgpool_hostname0□□□□□pgpool□□□□□□□IP□

·other_pgpool_port0□□□□□pgpool□□□□□□□

·other_wd_port0□□□□□pgpool□□□watchdog□□□□

□□□pgpool.conf□□□□□□□□□□□□□□□□□□□□□□□□□□□□□pgpool□□□□http://www.pgpool.net/docs/latest/en/html/runtime-config.html □

□□ pgpool□□□□pgpool.conf□pgpool□□□□□□□□□□□□□□□□□#□□□□□□pgpool□□□□□□□□□□□□□□□□pgpool□□□

/etc/pgpool-II/failover_stream.sh□□□□□□□□□

```
#! /bin/bash
# Execute command by failover.
# special values:  %d = node id
#                  %h = host name
#                  %p = port number
#                  %D = database cluster path
#                  %m = new master node id
#                  %M = old master node id
#                  %H = new master node host name
#                  %P = old primary node id
#                  %R = new master database cluster path
#                  %r = new master port number
#                  %% = '%' character

falling_node=$1          # %d
old_primary=$2           # %P
new_primary=$3           # %H
pgdata=$4                # %R
pghome=/opt/pgsql
log=/tmp/failover.log
date >> $log

# □□□□□□□□□□□□□□□□□□□□□
echo "falling_node=$falling_node" >> $log
echo "old_primary=$old_primary" >> $log
echo "new_primary=$new_primary" >> $log
echo "pgdata=$pgdata" >> $log

# □□□□□□□□□□□□□□□□□□□□□□□□□root
if [ $falling_node = $old_primary ] && [ $UID -eq 0 ]; then
# □□□□□□□□□□□□□□□□□□□□□$PGDATA□□□□□□recovery.conf□□□□□
    if [ -f $pgdata/recovery.conf ]; then
        su postgres -c "$pghome/bin/pg_ctl promote -D $pgdata"
        echo "Local promote" >> $log
    else
        su postgres -c "ssh -T postgres@$new_primary $pghome/bin/pg_ctl promote -D $pgdata"
        echo "Remote promote" >> $log
    fi
fi;
exit 0;
```

[illegible]pgpool[illegible]pgpool[illegible]$PGDATA[illegible]recovery.conf[illegible]pgpool[illegible]

[illegible]pgpool[illegible]ip addr[illegible]IP[illegible]root[illegible]pgpool[illegible]root[illegible]

```
export PGPOOL_HOME=/opt/pgpool
export PATH=$PGPOOL_HOME/bin:$PATH:.
```

[illegible]pghost4[illegible]root[illegible]pgpool[illegible]

```
# pgpool
```

[illegible]pgpool[illegible]pgpool[illegible]pgpool[illegible]

·-a[illegible]--hba-file[illegible]pgpool[illegible]hba[illegible]$prefix/etc/pool_hba.conf[illegible]

·-f[illegible]--config-file[illegible]pgpool[illegible]$prefix/etc/pgpool.conf[illegible]

·-F□--pcp-file□□□pgpool□pcp□□□□□□□$prefix/etc/pcp.conf□

□□□□□□□□□□/opt/pgpool/etc□□□□□□□pgpool□□□□□□□□□□□

□□□□□pgpool.conf□□□□□□□□□□□□□□□reload□□□□□□□□□□□□

```
# pgpool reload
```

□□pgpool□□□□□□□

```
# pgpool -m fast stop
```

pgpool□□□□□□□□□□□□□smart□fast□immediate□smart□□□□□□□□□□□□□□□□□pgpool□fast□immediate□□□□□□□□□□□□□pgpool□□□□□□□□□□□□□□□□fast□□□

□□pghost4□□□□□/var/log/messages□□□□□□□□□□□□□□□□□□□□□□□□□□pghost5□□pgpooll□□□□□□□pghost5□□/var/log/messages□□□□□□□□□□□□□pgpool□□□□□□□□

[illegible]VIP[illegible]pgpool[illegible]

```
[postgres@pghost4 ~]$ psql -h 192.168.26.72 -p 9999 postgres
postgres
postgres=# show pool_nodes;
    node_id | hostname | port | status | lb_weight |  role
| select_cnt | load_balance_node | replication_delay
---------+---------+------+-----+---------+-------+-------
---+----------------
    0     | pghost4  | 1921 | up   | -nan    | primary | 0
| true         | 0
    1     | pghost5  | 1921 | up   | -nan    | standby | 0
| false        | 0
(2 rows)
```

[illegible]pgpool[illegible]9999[illegible]pgpool[illegible]status[illegible]up[illegible]pgpool[illegible]down[illegible]pgpool[illegible]role[illegible]primary[illegible]standby[illegible]node_id[illegible]0[illegible]pghost4[illegible]node_id[illegible]1[illegible]pghost5[illegible]

[illegible] [illegible]

14.1.3 PCP[illegible]

pgpool[illegible]pgpool[illegible]pgpool[illegible]pgpool[illegible]pgpool[illegible]

PCP□□□□□□□□□pgpool□□□□PostgreSQL□□□□□□□□□□□□□pcp□□□□□□□MD5□□□□□□□□□□pcp.conf□□□□□□pcp.conf□□□□□□□□□

```
# USERID:MD5PASSWD
```

USERID□□□□□□MD5PASSWD□□□□□□□□□□□□□□□□□□

□□□□□□□□pgpool□□□□□PCP□□□□□□□□□PCP□□□□□□□□pgpool□□□pg_md5□□MD5□□□□□□□□

```
# pg_md5 pgpool
ba777e4c2f15c11ea8ac3be7e0440aa0
```

□□pcp.conf□□□□□□□

```
# USERID:MD5PASSWD
pgpool:ba777e4c2f15c11ea8ac3be7e0440aa0
```

PCP□□□□□pcp_node_info□pcp_watchdog_info□pcp_attach_node□□□□□□pcp_node_info□□□□□□□□□□□□□□

```
$ pcp_node_info --verbose -h 192.168.26.72 -U pgpool 0
Hostname   : pghost4
Port       : 1921
Status     : 2
Weight     : nan
Status Name: up
```

[illegible]0[illegible]Status[illegible]

·0[illegible]PCP[illegible]

·1[illegible]

·2[illegible]

·3[illegible]

[illegible]pcp_attach_node[illegible]

14.1.4 pgpool[illegible]

[illegible]pgpool[illegible]

·[illegible]pgpool[illegible]pgpool[illegible]

·[illegible]

·[illegible]

[illegible]pgpool[illegible]pgpool[illegible]

pghost4[illegible]pgpool[illegible]pghost5[illegible]pgpool[illegible]VIP[illegible]pghost4[illegible]

[illegible]pcp_watchdog_info[illegible]pgpool[illegible]watchdog[illegible]watchdog[illegible]pgpool[illegible]watchdog[illegible]pgpool[illegible]

```
[postgres@pghost4 ~]$ pcp_watchdog_info --verbose -h
192.168.26.72 -U pgpool
Password:
Watchdog Cluster Information
Total Nodes          : 2
Remote Nodes         : 1
Quorum state         : QUORUM EXIST
Alive Remote Nodes   : 1
VIP up on local node : YES
Master Node Name     : pghost4:9999 Linux pghost4
Master Host Name     : pghost4

Watchdog Node Information
Node Name      : pghost4:9999 Linux pghost4
Host Name      : pghost4
Delegate IP    : 192.168.26.72
Pgpool port    : 9999
Watchdog port  : 9000
Node priority  : 1
Status         : 4
Status Name    : MASTER

Node Name      : pghost5:9999 Linux pghost5
```

```
Host Name       : pghost5
Delegate IP     : 192.168.26.72
Pgpool port     : 9999
Watchdog port   : 9000
Node priority   : 1
Status          : 7
Status Name     : STANDBY
```

□□□□□watchdog□□□□□pghost4□watchdog□□□□pghost5□watchdog□□□□□pgpool□□□□□□watchdog□□□□□□□□□□

□□pghost4□□IP□□□□□□□□□□

```
[postgres@pghost4 ~]$ ip a
...
12: bond0: <BROADCAST,MULTICAST,MASTER,UP,LOWER_UP> mtu 1500
qdisc noqueue state UP
    link/ether a0:36:9f:9b:07:af brd ff:ff:ff:ff:ff:ff
    inet 192.168.26.57/24 brd 192.168.26.255 scope global
bond0
    inet 192.168.26.72/24 scope global secondary bond0:1
...
```

□□192.168.26.72 VIP□pghost4□□□□□□pghost4□pgpool□□□□

pghost4□□□pgpool□□□□□□□□□

```
# pgpool -m fast stop
.done.
```

□□pgpool□□/var/log/messages□□□□□□□□□□

```
Oct 12 10:09:39 pghost4 pgpool[41818]: [13-1] 2017-10-12
10:09:39: pid 41818: LOG:  received fast shutdown request
Oct 12 10:09:39 pghost4 pgpool[41818]: [14-1] 2017-10-12
10:09:39: pid 41818: LOG:  shutdown request. closing listen
socket
Oct 12 10:09:39 pghost4 pgpool[26337]: [1-1] 2017-10-12
10:09:39: pid 26337: LOG:  stop request sent to pgpool.
waiting for termination...
Oct 12 10:09:39 pghost4 pgpool: watchdog[41819]: [45-1]
2017-10-12 10:09:39: pid 41819: LOG:  Watchdog is shutting
down
Oct 12 10:09:39 pghost4 pgpool: watchdog de-
escalation[26338]: [46-1] 2017-10-12 10:09:39: pid 26338:
LOG:  watchdog: de-escalation started
Oct 12 10:09:39 pghost4 pgpool: watchdog de-
escalation[26338]: [47-1] 2017-10-12 10:09:39: pid 26338:
LOG:  successfully released the delegate IP:"192.168.26.72"
Oct 12 10:09:39 pghost4 pgpool: watchdog de-
escalation[26338]: [47-2] 2017-10-12 10:09:39: pid 26338:
DETAIL:  'if_down_cmd' returned with success
Oct 12 10:09:41 pghost4 ntpd[6835]: Deleting interface #14
bond0:1, 192.168.26.72#123, interface stats: received=0,
sent=0, dropped=0, active_time=390 secs
```

□□□□□□□□□□□pgpool□□□□fast□□□□□□□□□□□□□□□□□□□□watchdog□□□□□pgpool.conf□□□□□if_down_cmd□□□□□□□□□□VIP□Deleting interface□□□□VIP□□□□□

pghost4□□□□□IP□□□□□□□VIP□□□□□□□□□□□□

```
[postgres@pghost4 ~]$ ip a
12: bond0: <BROADCAST,MULTICAST,MASTER,UP,LOWER_UP> mtu 1500
qdisc noqueue state UP
    link/ether a0:36:9f:9b:07:af brd ff:ff:ff:ff:ff:ff
    inet 192.168.26.57/24 brd 192.168.26.255 scope global
bond0
```

□□□□pghost5□IP□□□□□VIP□□□□□pghost5□□□bond0□□□

□□□□pcp_watchdog_info□□□□pgpool□□□□□pghost5□pgpool□□□□□MASTER□□□pghost4□□pgpool□□□SHUTDOWN□

□□pgpool□□□□□□□□□□□□□□□□

```
[postgres@pghost4 ~]$ psql -h 192.168.26.72 -p 9999 postgres
postgres
Password for user postgres:
psql (10.0)
Type "help" for help.
```

pgpool□□□□□□□□□□□□

□□□□□□□□□□□□□pgpool□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□pghost4□□□□□□pghost5□□□□□□□□□□□□pgpool□□□□□□□□□□□□□□□□□□pgpool□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□pghost4□pgpool□□□□□□□□□□□□□pghost5□pgpool□□□□□□□□□□□□□□□□VIP□pghost4□□

□□pgpool□□□□□□□□□□□□□□□□□

```
[postgres@pghost4 ~]$ psql -h 192.168.26.72 -p 9999 postgres
postgres
postgres=# show pool_nodes;
    node_id | hostname | port | status | lb_weight |  role
| select_cnt | load_balance_node | replication_delay
------------+--------+-----+-----+-------+-------+--------
-+----------------
    0       | pghost4 | 1921 | up   | -nan    | primary | 0
| true        | 0
    1       | pghost5 | 1921 | up   | -nan    | standby | 0
| false       | 0
(2 rows)
```

pghost4□□□□□□□□□□□

```
[postgres@pghost4 ~]$ pg_ctl stop -m fast
waiting for server to shut down.... done
server stopped
```

□□□□□□□□□pgpool□□□□□pgpool.conf□□□□□failover_command□□□□□/opt/pgpool/failover_stream.sh□□□□□□□□□□□□□/tmp/failover.log□□□□□□□□□□□□□□□□

```
Thu Oct 12 11:01:50 CST 2017
falling_node=0
old_primary=0
new_primary=pghost5
pgdata=/data1/pg10/pg_root
Remote promote
```

falling_node□□□□PostgreSQL□□□□□□old_primary□□□□□□□□□new_primary□□□□□□□□□□pgdata□□□□□□□□□□

pghost5□□□□□□□□□□□□□□

```
[postgres@pghost5 ~]$ pg_controldata | grep cluster
Database cluster state:               in production
```

□□pghost5□□□□□□□□□□□□□□□pghost5□□□□□□□□□□□□□□□□□

□□□□pgpool□□□□□□□□□□□□□□□

```
[postgres@pghost4 ~]$ psql -h 192.168.26.72 -p 9999 postgres
postgres
postgres=# show pool_nodes;
    node_id | hostname | port | status | lb_weight |  role
| select_cnt | load_balance_node | replication_delay
---------+----------+------+-------+----------+---------+-
-------+------------
    0    | pghost4  | 1921 | down   | -nan      | standby |
0      | false    | 0
    1    | pghost5  | 1921 | up     | -nan      | primary |
```

```
0      | true     | 0
(2 rows)
```

□□pghost4□□□□□□down□□□□□□□□standby□pghost5□□□□□□□□up□□□□□□□□primary□

□□□□pgpool□□□□□pgpool□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□pgpool□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□pghost4□pgpool□□□□□□□□□□□□□pghost5□pgpool□□□□□□□□□□□□□□□VIP□pghost4□□

□□pghost4□□□□□pghost4□□□□pgpool□□□□□□□□□□□□□□□□pghost5□□pgpool□□□□□□□□□□□□□□□□□□□□

```
[root@pghost4 pgpool]# reboot
```

□□pghost5□□/tmp/failover.log□□□□□□□□

```
Thu Oct 12 11:43:13 CST 2017
falling_node=0
```

```
old_primary=0
new_primary=pghost5
pgdata=/data1/pg10/pg_root
Local promote
```

□□□□□□pghost4□□□□pghost5□□□/etc/pgpool-II/failover_stream.sh□□□

pghost5□□□□□□□□□□□□□□□□

```
[postgres@pghost5 ~]$ pg_controldata | grep cluster
Database cluster state:               in production
```

□pghost5□□□□□□□□□□□

□pgpool□□□□□□□□

```
[postgres@pghost5 ~]$ psql -h 192.168.26.72 -p 9999 postgres postgres
postgres=# show pool_nodes;
    node_id | hostname | port | status | lb_weight |  role 
| select_cnt | load_balance_node | replication_delay
------------+----------+------+--------+-----------+---------
-+-----+------------
    0       | pghost4   | 1921  | down    | -nan       | 
standby  | 0    | false   |   0
    1       | pghost5   | 1921  | up      | -nan       | 
primary  | 0    | true    |   0
(2 rows)
```

□□□phost4□□□□down□□□□□standby□pghost5□□□□□up□□□□□

primary[illegible]

[illegible]pghost5[illegible]/var/log/messages[illegible]pgpool[illegible]

```
...
Oct 12 11:43:15 pghost5 ntpd[6835]: Listen normally on 9
bond0:1 192.168.26.72 UDP 123
Oct 12 11:43:17 pghost5 pgpool: watchdog escalation[44067]:
[125-1] 2017-10-12 11:43:17: pid 44067: LOG:  successfully
acquired the delegate IP:"192.168.26.72"
Oct 12 11:43:17 pghost5 pgpool: watchdog escalation[44067]:
[125-2] 2017-10-12 11:43:17: pid 44067: DETAIL:  'if_up_cmd'
returned with success
Oct 12 11:43:17 pghost5 pgpool: watchdog[42054]: [145-1]
2017-10-12 11:43:17: pid 42054: LOG:  watchdog escalation
process with pid: 44067 exit with SUCCESS.
...
```

[illegible]pgpool.conf[illegible]if_up_cmd[illegible]VIP[illegible]ip addr[illegible]

```
[postgres@pghost5 ~]$ ip a
...
12: bond0: <BROADCAST,MULTICAST,MASTER,UP,LOWER_UP> mtu 1500
qdisc noqueue state UP
    link/ether a0:36:9f:9d:c7:5f brd ff:ff:ff:ff:ff:ff
    inet 192.168.26.58/24 brd 192.168.26.255 scope global
bond0
    inet 192.168.26.72/24 scope global secondary bond0:1
...
```

[illegible]VIP[illegible]pghost5[illegible]

□□□□□□□□□□□□□□□□12.5□□

□□□□□□□□□□□□□□□□□□□□□□□

·□□□pgpool□□□□□□□□□□□□□□□□□pgpool□
□□□□□□□□□□□□□□

·□□□PostgreSQL□□□□□□□□□□□pgpool□□□
□□PostgreSQL□□□□□□□□□□□□□□

·□□□□□□□□□□□□□□pgpool□PostgreSQL□□
□□□□□□□□□□pgpool□VIP□□□pgpool□□□□

□□□□□□□pgpool□□□□□□□□□□□□□□□□□□□□
□□□□

□□ □□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□pgpool□□□□□□
□□□failover_stream.sh□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□Keepalived+□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□

14.1.5 pgpool□□□□□□□□

□□□□□pgpool□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□pgpool□□□□□□□□□□□□□□

□□□□□□pgpool□□□□pool_passwd□□□□□□
□□□

```
[postgres@pghost4 ~]$ psql -h 192.168.26.72 -p 9999 postgres
pguser
psql: FATAL:  md5 authentication failed
DETAIL:  pool_passwd file does not contain an entry for
"pguser"
```

□□□□□pgpool□enable_pool_hba□□□
pgpool□□□pool_hba.conf□□□□□□□□□
pool_hba.conf□□□□□□□md5□□□□□□□□
pool_passwd□□□□□□□□□□□14.1.2□□□□□□
pool_passwd□□□□□□□repuser□□□□□□
pool_passwd□□□□□□□□□□□□□□□

```
repuser:md54f87427f75b5a59ba0abffe11a6f79a8
```

□□□□□□□pgpool□□□□□□□□□□□□□□□□□□□□
□□

□□□□□□pcp□□□□□□□□□□□

□□□□□□□□□

```
[postgres@pghost5 ~]$ pcp_node_info -h 192.168.26.72 -U
pgpool 0
Password:
```

```
FATAL:  authentication failed for user "pgpool"
DETAIL:  username and/or password does not match
```

14.1.3[illegible]pcp[illegible]pcp.conf[illegible]pcp.conf[illegible]pcp.conf[illegible]

```
# USERID:MD5PASSWD
pgpool:ba777e4c2f15c11ea8ac3be7e0440aa0
```

[illegible]pgpool[illegible]VIP[illegible]pgpool[illegible]

pgpool[illegible]VIP[illegible]pgpool.conf[illegible]if_up_cmd[illegible]if_down_cmd[illegible]VIP[illegible]

```
if_up_cmd = 'ip addr add $_IP_$/24 dev bond0 label bond0:1'
if_down_cmd = 'ip addr del $_IP_$/24 dev bond0'
```

[illegible]bond0[illegible]VIP[illegible]bond0[illegible]1[illegible]

[illegible]pgpool[illegible]VIP[illegible]pgpool[illegible]pgpool[illegible]if_up_cmd[illegible]if_down_cmd[illegible]ip addr add[illegible]ip addr del[illegible]

```
$ ip addr add 192.168.26.72/24 dev bond0 label bond0:1
$ ip addr del 192.168.26.72/24 dev bond0:1
```

[illegible]ip addr[illegible]if_up_cmd[illegible]if_down_cmd[illegible]

[illegible]

[illegible]pgpool[illegible]pgpool.conf[illegible]failover_command[illegible]'/opt/pgpool/failover_stream.sh%d%P%H %R'[illegible]

[illegible]failover_stream.sh[illegible]/tmp/failover.log[illegible]failover_stream.sh[illegible]

[illegible]/tmp/failover.log[illegible]

```
Thu Oct 12 11:21:20 CST 2017
falling_node=1
old_primary=1
new_primary=pghost4
pgdata=/data1/pg10/pg_root
Local promote
```

□□/failover_stream.sh□□□□□□□□falling_node□□□old_primary□□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□PostgreSQL□□□□□□□□□□pcp_attach_node□□□□□□□□□□pgpool□

pgpool□□□□□□□□□□□

```
postgres=# show pool_nodes;
    node_id | hostname | port | status | lb_weight |  role
| select_cnt | load_balance_node | replication_delay
------------+----------+------+--------+-----------+--------
-+------------+----
    0       | pghost4  | 1921 | up     | -nan      | primary
| 0   | true    | 0
    1       | pghost5  | 1921 | up     | -nan      | standby
| 0   | false   | 0
(2 rows)
```

□pghost5□□□□□□□□□□□□

```
[postgres@pghost5 ~]$ pg_ctl stop -m fast
waiting for server to shut down.... done
server stopped
```

□□□□□□pgpool□□show pool_nodes□□□□□□□□□□□pghost5□□□□□□down□□□pghost5□□□□□□□□□□□

```
[postgres@pghost5 ~]$ pg_ctl start
```

□□□□pgool□□□□□□□□□□

```
postgres=# show pool_nodes;
    node_id | hostname | port | status | lb_weight |  role
| select_cnt | load_balance_node | replication_delay
----------+--------+-----+------+---------+--------+------
----+---------------
    0       | pghost4  | 1921 | up     | -nan      | primary
| 0          | true         | 0
    1       | pghost5  | 1921 | down   | -nan      | standby
| 0          | false        | 0
(2 rows)
```

pghost5□□□□□□□□down□□□□□□□
pcp_attach_node□□□PostgreSQL□□□□□□
pgpool□□□□□□□□

```
[postgres@pghost4 ~]$ pcp_attach_node -h 192.168.26.72 -U
pgpool 1
Password:
pcp_attach_node -- Command Successful
```

pcp_attach_node□□□□□□□□□

```
pcp_attach_node [options...] [node_id]
```

node_id□PostgreSQL□□□□□□pghost4□□□□□□□0□pghost5□□□□□□□1□□□□pghost5□□□□□□□□□pgpool□

□□□□pgpool□□show pool_nodes□□□□□□□□□□□pghost5□□□□□□up□

14.2 □□Keepalived+□□□□□□□□□

□□□□□□□□□□pgpool□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□Keepalived□□□□□□□□□□□□□□□□□□□□Keepalived□□□□□□□□□PostgreSQL□□□□□□□□□□Keepalived□□□□□□□□□□Keepalived□□□□□□VIP□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□

·□□□□□□□□□□□□□□□□PostgreSQL□□□□□□□

·□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

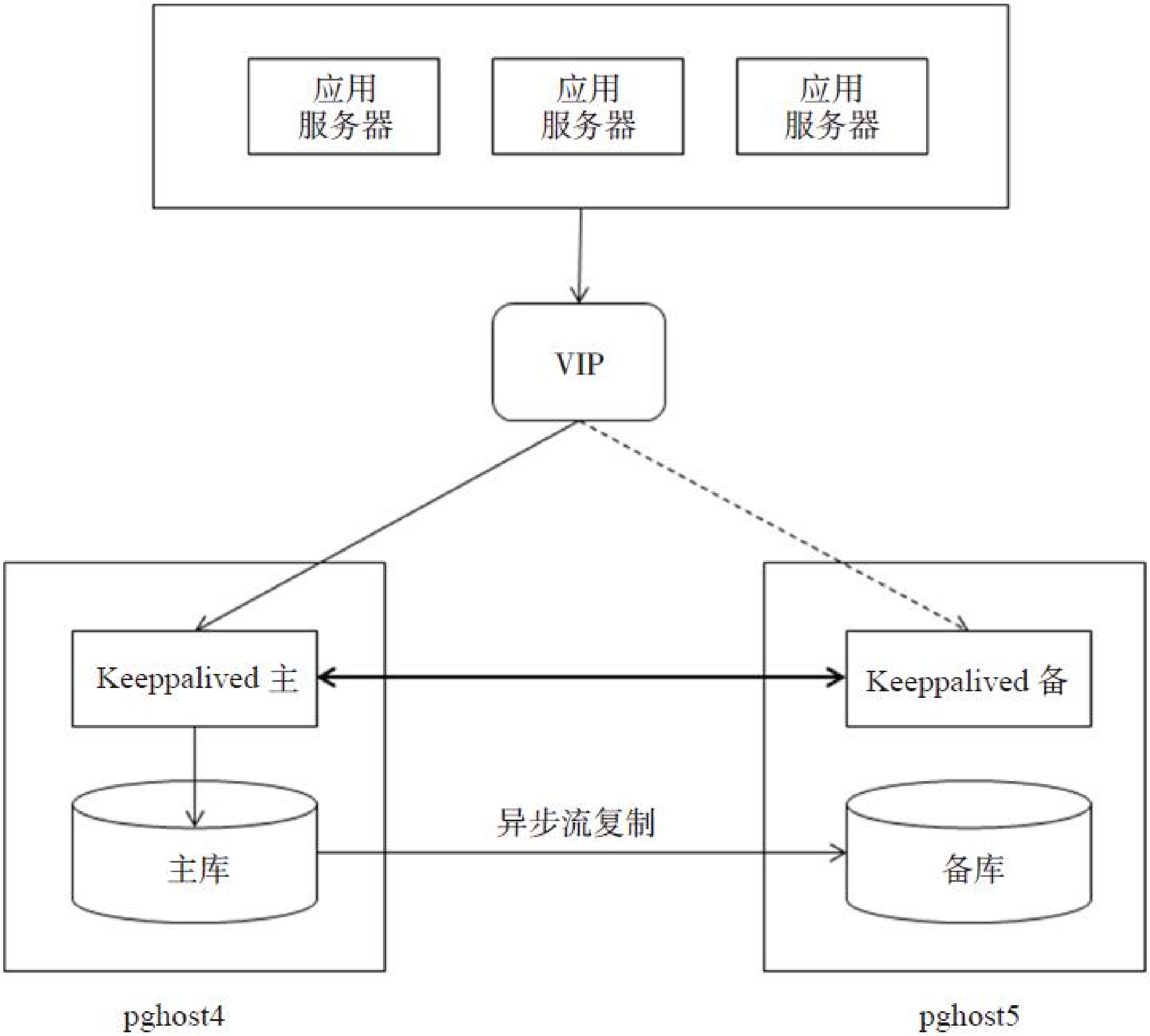
□□□□□PostgreSQL10□□□□□□□□□14-2□

□14-2　Keepalived+□□□□□□□□□□

主机名	组件	IP 地址	版本	VIP
pghost4	主库	192.168.26.57	PostgreSQL10	192.168.26.72
	Keepalived 主	192.168.26.57	Keepalived-1.3.7	
pghost5	备库	192.168.26.58	PostgreSQL10	
	Keepalived 备	192.168.26.58	Keepalived-1.3.7	

14.2.1 Keepalived+□□□□□□□□□□□

Keepalived+□□□□□□□□□□□□□14-2□



□14-2 Keepalived+□□□□□□□□□□

Keepalived□□□□□□□□□□□□pghost4□□□□Keepalived□□□□□□□□□□□□pghost5□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

[illegible]Keepalived[illegible][illegible]Keepalived[illegible][illegible]Keepalived[illegible]VIP[illegible]Keepalived[illegible]

14.2.2 Keepalived+[illegible]

[illegible]

1[illegible]pghost4[illegible]pghost5[illegible]

2[illegible]Keepalived[illegible]Keepalived[illegible]sr_delay[illegible]Keepalived[illegible]last_alive[illegible]

```
postgres=# CREATE ROLE keepalived NOSUPERUSER NOCREATEDB
           login ENCRYPTED PASSWORD 'keeplaived';
CREATE ROLE
postgres=# CREATE DATABASE keepalived
           WITH OWNER=keepalived
           TEMPLATE=TEMPLATE0
           ENCODING='UTF8';
CREATE DATABASE

postgres=# \c keepalived keepalived
keepalived=> CREATE TABLE sr_delay(id int4,last_alive
timestamp(0) without time zone);
CREATE TABLE
```

[illegible]sr_delay[illegible]

```
CREATE FUNCTION cannt_delete ()
RETURNS trigger
LANGUAGE plpgsql AS $$
BEGIN
RAISE EXCEPTION 'You can not delete!';
END; $$;
```

[illegible]cannt_delete[illegible]cannt_truncate[illegible]

```
keepalived=> CREATE TRIGGER cannt_delete BEFORE DELETE ON
sr_delay
    FOR EACH ROW EXECUTE PROCEDURE cannt_delete();
CREATE TRIGGER
keepalived=> CREATE TRIGGER cannt_truncate BEFORE TRUNCATE
ON sr_delay
    FOR STATEMENT EXECUTE PROCEDURE cannt_delete();
CREATE TRIGGER
```

sr_delay[illegible]

```
keepalived=> INSERT INTO sr_delay VALUES(1,now());
INSERT 0 1
```

[illegible]Keepalived[illegible]PostgreSQL[illegible]Keepalived[illegible]Keepalived[illegible]pg_hba.conf[illegible]

```
# keepalived
host keepalived keepalived 192.168.26.57/32 md5
host keepalived keepalived 192.168.26.58/32 md5
host keepalived keepalived 192.168.26.72/32 md5
```

□□□□pg_ctl reload□□□□□□□□

3□pghost4□pghost□□Keepalived□□□

□□Keepalived□□□□□□□□□□□
http://www.keepalived.org/software/keepalived-1.3.7.tar.gz □

□□□□□□□□□□□□□

```
# yum install openssl openssl-devel popt popt-devel
```

pghost4□pghost□□□□□□□Keepalived□□□□□□

```
# tar xvf keepalived-1.3.7.tar.gz
# ./configure --prefix=/usr/local/keepalived
# make
# make install
```

□Keepalived□□□□□□□□□□□□□□□□

```
# ln -s /usr/local/keepalived/sbin/keepalived /usr/sbin/
# cp /opt/soft_bak/keepalived-
1.3.7/keepalived/etc/init.d/keepalived /etc/init.d/
# cp /usr/local/keepalived/etc/sysconfig/keepalived
/etc/sysconfig
```

□□□□□□□□□□□□□□□□□□□□□Keepalived□□□□□□□/etc/init.d□□□□□□□□□□□□□□□□□□□□□□□

```
# service keepalived status
keepalived is stopped
```

□□□□Keepalived□□□□□□□□Keepalived□□□□□□□□□□□□□□□

```
# chkconfig keepalived on
```

14.2.3 Keepalived□□

□□□□□□□Keepalived□□□□□□□□□□□□□□□□Keepalived□

□□Keepalived□□□□□□□□□□

```
# mkdir -p /etc/keepalived
```

□□/etc/keepalived/keepalived.conf□□□□□□□□□□□□

```
! Configuration File for keepalived

global_defs {
    notification_email {
        francs3@163.com
    }
    smtp_server 127.0.0.1
    smtp_connect_timeout 30
    router_id DB1_PG_HA
}

vrrp_script check_pg_alived {
    script "/usr/local/bin/pg_monitor.sh"
    interval 10
    fall 3    # require 3 failures for KO
}

vrrp_instance VI_1 {
    state BACKUP
    nopreempt
    interface bond0
    virtual_router_id 10
    priority 100
    advert_int 1
    authentication {
        auth_type PASS
        auth_pass t9rveMP0Z9S1
    }
        track_script {
        check_pg_alived
    }
    virtual_ipaddress {
        192.168.26.72
    }
    smtp_alert
    notify_master /usr/local/bin/active_standby.sh
}
```

□□Keepalived□□□□□Keepalived□□□
priority□□□90□□□□□□□□□□□□□□□□□

·global_defs□□□□□□□□□□
Keepalived□□□□□□□□□□□□

·vrrp_script□□□□□□□□□10□□□□
□/usr/local/bin/pg_monitor.sh□□□□□□□□
fall 3□□□□□□□□□

·vrrp_instance□vrrp□□□□□□□□□□□□
□□□ID□□□□□□□backup□□□□□□□
nopreempt□□□□□□□□VIP□□□□priority□□
□□□□□□□□□□□□Keepalived□priority□□□
□Keepalived□□□□□□□Keepalived□VIP□□□
□□□□bond0□

·smtp_alert□□□notify_master□□□
Keepalived□□□□□□□□□
□/usr/local/bin/active_standby.sh□□□□□□
□□□

/usr/local/bin/pg_monitor.sh□□□□□□

```
#!/bin/bash
# □□□□
export PGPORT=1921
export PGUSER=keepalived
```

```
export PGDBNAME=keepalived
export PGDATA=/data1/pg10/pg_root
export LANG=en_US.utf8
export PGHOME=/opt/pgsql
export
LD_LIBRARY_PATH=$PGHOME/lib:/lib64:/usr/lib64:/usr/local/lib
64:/lib:/usr/lib:/usr/local/lib
export PATH=$PGHOME/bin:$PGPOOL_HOME/bin:$PATH:.

MONITOR_LOG="/tmp/pg_monitor.log"
SQL1="UPDATE sr_delay SET last_alive = now();"
SQL2='SELECT 1;'
# □□□□□□□□□□□□□□□□□□□□□
standby_flg=`psql -p $PGPORT -U postgres -At -c "SELECT
pg_is_in_recovery();"`
if [ ${standby_flg} == 't' ]; then
    echo -e "`date +%F\ %T`: This is a standby database,
exit!\n" >> $MONITOR_LOG
    exit 0
fi

# □□□□sr_delay□
echo $SQL1 | psql -At -p $PGPORT -U $PGUSER -d $PGDBNAME >>
$MONITOR_LOG

# □□□□□□□□
echo $SQL2 | psql -At -h -p $PGPORT -U $PGUSER -d $PGDBNAME
if [ $? -eq 0 ]; then
    echo -e "`date +%F\ %T`:  Primary db is health."  >>
$MONITOR_LOG
    exit 0
else
    echo -e "`date +%F\ %T`:  Attention: Primary db is not
health!" >> $MONITOR_LOG
    exit 1
fi
```

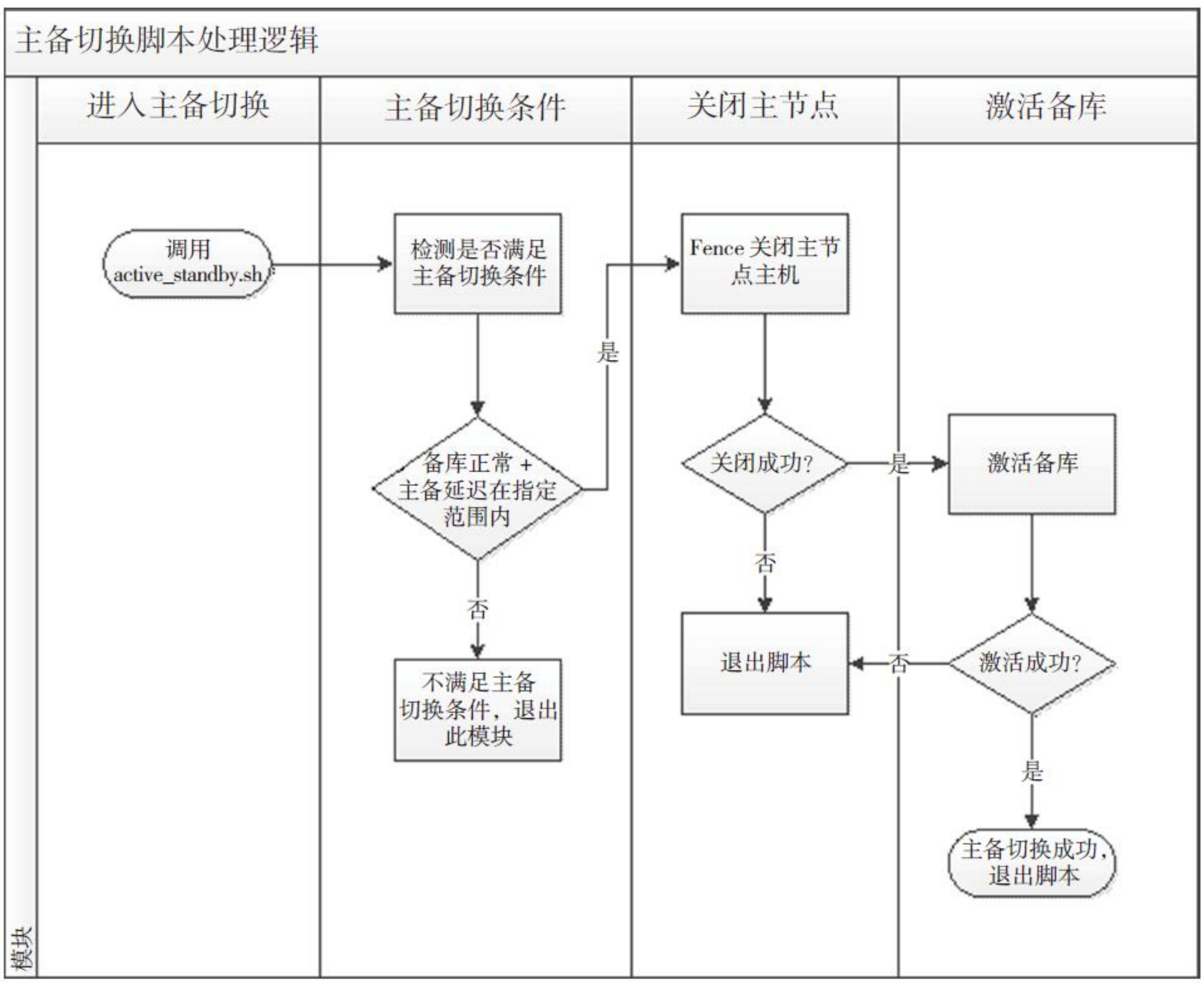
□□□□□10□□□□□□□□□□□keepalived.conf□□□□□interval□□□□□□□□□□□□□

·□□□□□□□□□

·□□sr_delay□last_alive□□□□□□□□□

□Keepalived□□□□□□□□□□□□Keepalived□□□□□□□Keepalived□□□□□□□□□□□□notify_master□□□□□/usr/local/bin/active_standby.sh□□□

/usr/local/bin/active_standby.sh□□□□□□□□14-3□□□



□14-3 □□□□□□active_stadnby.sh□□□□

□□□□□□□□active_standby.sh□□□□□□□□□□
□□□

1□□□□□□□□□□□□□□□

2□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□

1□□□□□□□□□□□□□□□

2□□□□□□

□□□

/usr/local/bin/active_standby.sh□□□□□□
□□□

```
#/bin/bash
# □□□□
export PGPORT=1921
export PGUSER=keepalived
export PG_OS_USER=postgres
export PGDBNAME=keepalived
export PGDATA=/data1/pg10/pg_root
export LANG=en_US.utf8
export PGHOME=/opt/pgsql
export
LD_LIBRARY_PATH=$PGHOME/lib:/lib64:/usr/lib64:/usr/local/lib
64:/lib:/usr/lib:/usr/local/lib
```

```
export
PATH=/opt/pgbouncer/bin:$PGHOME/bin:$PGPOOL_HOME/bin:$PATH:.

# □□□□, LAG_MINUTES □□□□□□□□□□□□□□
LAG_MINUTES=60
HOST_IP=`hostname -i`
NOTICE_EMAIL="francs3@163.com"
FAILOVE_LOG='/tmp/pg_failover.log'

SQL1="SELECT 'this_is_standby' AS cluster_role FROM ( SELECT
pg_is_in_recovery() AS std ) t WHERE t.std is true;"
SQL2="SELECT 'standby_in_allowed_lag' AS cluster_lag FROM
sr_delay WHERE now()-last_alive < interval '$LAG_MINUTES
SECONDS';"

# □□□□□□□□□□IP□□□□□□□
FENCE_IP=50.1.225.101
FENCE_USER=root
FENCE_PWD=xxxx

# VIP □□□□□□□□□□□□□
echo -e "`date +%F\ %T`: keepalived VIP switchover!" >>
$FAILOVE_LOG

# VIP □□□□□□□□
#echo -e "`date +%F\ %T`: ${HOST_IP}/${PGPORT} VIP □□□□□□□□□
□□\n\nAuthor: francs(DBA)" | mutt -s "Error: □□□ VIP □□□□ "
${NOTICE_EMAIL}

# pg_failover □□□□□□□□□□□□□
pg_failover()
{
# FENCE_STATUS  □□□□□□□□□□□□□□□□□□1 □□□□□0 □□□□
# PROMOTE_STATUS  □□□□□□□□□□□1 □□□□□0 □□□□
FENCE_STATUS=1
PROMOTE_STATUS=1

# □□□□□□□□□□□□□□□□□□□
for ((k=0;k<10;k++))
do
# □□ipmitool□□□□□□□□□□□□□□□□□□X86□□□□□□□□□
    ipmitool -I lanplus -L OPERATOR -H $FENCE_IP -U
$FENCE_USER -P $FENCE_PWD power reset
    if [ $? -eq 0 ]; then
        echo -e "`date +%F\ %T`: fence primary db host
```

```
success."
        FENCE_STATUS=0
        break
    fi
sleep 1
done

if [ $FENCE_STATUS -ne 0 ]; then
    echo -e "`date +%F\ %T`: fence failed. Standby will not
promote, please fix it manually."
return $FENCE_STATUS
fi

# □□□□
su - $PG_OS_USER -c "pg_ctl promote"
if [ $? -eq 0 ]; then
    echo -e "`date +%F\ %T`: `hostname` promote standby
success. "
    PROMOTE_STATUS=0
fi

if [ $PROMOTE_STATUS -ne 0 ]; then
    echo -e "`date +%F\ %T`: promote standby failed."
    return $PROMOTE_STATUS
fi

    echo -e "`date +%F\ %T`: pg_failover() function call
success."
    return 0
}

# □□□□□□
# □□□□□□□□□□STANDBY_CNT=1 □□□□.
STANDBY_CNT=`echo $SQL1 | psql -At -p $PGPORT -U $PGUSER -d
$PGDBNAME -f - | grep -c this_is_standby`
echo -e "STANDBY_CNT: $STANDBY_CNT"  >> $FAILOVE_LOG

if [ $STANDBY_CNT -ne 1 ]; then
    echo -e "`date +%F\ %T`: `hostname` is not standby
database□failover not allowed! " >> $FAILOVE_LOG
    exit 1
fi

# □□□□□□□□□□□□□□, LAG=1 □□□□□□□□□□□□□
LAG=`echo $SQL2 | psql -At -p $PGPORT -U $PGUSER -d
```

```
$PGDBNAME | grep -c standby_in_allowed_lag`
echo -e "LAG: $LAG"  >> $FAILOVE_LOG

if [ $LAG -ne 1 ]; then
    echo -e "`date +%F\ %T`: `hostname` is laged far
$LAG_MINUTES SECONDS from primary , failover not allowed! "
>> $FAILOVE_LOG
    exit 1
fi

# [illegible]1[illegible]2[illegible]
if [ $STANDBY_CNT -eq 1 ] && [ $LAG -eq 1 ]; then
    pg_failover >> $FAILOVE_LOG
    if [ $? -ne 0 ]; then
        echo -e "`date +%F\ %T`: pg_failover failed." >>
$FAILOVE_LOG
        exit 1
    fi
fi

# [illegible]pg_failover[illegible]
# 1. [illegible]
# 2. [illegible]

# pg_failover[illegible]
# 1. [illegible]
# 2. [illegible]
```

[illegible]FENCE_IP[illegible]FENCE_USER[illegible]FENCE_PWD[illegible]

[illegible]

```
# chmod 700 /usr/local/bin/pg_monitor.sh
# chmod 700 /usr/local/bin/active_standby.sh
```

□□ □□□□□□□□□□□□□□□□□□□□□□□□□□
□□60□□□pg_monitor.sh□active_standby.sh□□□□□□□□□□□□

14.2.4 Keepalived□□□□□□□

□□□□□Keepalived+□□□

·□□□□□□Keepalived□□□□□□keepalived□□□□□□□□□□□□□□□□

·□□□□□□□□□□□□□□□□□□□□□□□□□□

·□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□Keepalived□□□□□□Keepalived□□□□□□□□□□□□□□□□

pghost4□□□Keepalived□□□□□□□□□□pghost5□□□Keepalived□□□□□□□□□□VIP□□□pghost4□□

pghost4□□□Keepalived□□□□□□Keepalived□□□□□□□□

```
[root@pghost4 ~]# ps -ef | grep keepalived | grep -v grep
root      7527     1  0 10:31 ?        00:00:00 keepalived -
D
root      7528  7527  0 10:31 ?        00:00:00 keepalived -
D
root      7529  7527  0 10:31 ?        00:00:00 keepalived -
D
    [root@pghost4 ~]# kill 7527
```

pghost5□□/var/log/messages□□□□□□□□□□

```
Oct 16 11:00:55 pghost5 Keepalived_vrrp[6561]:
VRRP_Instance(VI_1) Transition to MASTER STATE
Oct 16 11:00:56 pghost5 Keepalived_vrrp[6561]:
VRRP_Instance(VI_1) Entering MASTER STATE
Oct 16 11:00:56 pghost5 Keepalived_vrrp[6561]:
VRRP_Instance(VI_1) setting protocol VIPs.
...
Oct 16 11:00:57 pghost5 ntpd[6917]: Listen normally on 6
bond0 192.168.26.72 UDP 123
Oct 16 11:01:01 pghost5 Keepalived_vrrp[6561]: Sending
gratuitous ARP on bond0 for 192.168.26.72
...
```

□□□□□pghost5□□Keepalived□□□□□□□□□□□□□□□VIP 192.168.26.72□

pghost5□□□□□ip addr□□□□□□□□□□

```
[root@pghost5 ~]# ip addr
...
12: bond0: <BROADCAST,MULTICAST,MASTER,UP,LOWER_UP> mtu 1500
qdisc noqueue state UP
    link/ether a0:36:9f:9d:c7:5f brd ff:ff:ff:ff:ff:ff
    inet 192.168.26.58/24 brd 192.168.26.255 scope global
bond0
    inet 192.168.26.72/32 scope global bond0
...
```

□□VIP□□□□□pghost5□□□□

pghost5□□□□□/tmp/pg_failover.log□□□□□□□□□□□□□□active_standby.sh□□□□□□□□□□□

```
2017-10-16 11:00:56: keepalived VIP switchover!
STANDBY_CNT: 1
LAG: 1
Chassis Power Control: Reset
2017-10-16 11:00:56: fence primary db host success.
waiting for server to promote.... done
server promoted
2017-10-16 11:00:56: pghost5 promote standby success.
2017-10-16 11:00:56: pg_failover() function call success.
```

□□□□□□□□Keepalived□VIP□□□□□□□□□□□□□□□□□□□pghost5□□□□□□□□□□□□□pghost4□□□□□□□□

□□pghost5□□□□□□□□□□□□

```
[postgres@pghost5 pg_root]$ pg_controldata | grep cluster
Database cluster state:               in production
```

□□pghost5□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□pghost4□Keepalived□□□□□□□□□□□□□pghost5□Keepalived□□□□□□□□□□□□□□□VIP□pghost4□□□pghost4□□□□□□□□□□□□□□□□□□□□$PGDATA□□□□recovery.done□□□recovery.conf□□□□□□□□□

□□□□□□□□□□□□□□□□12.5□□

□□□□□□□□□□□□□□□□□□□□□□□□□□

pghost4□□□□□□□□□□

```
[postgres@pghost4 ~]$ pg_ctl stop -m fast
waiting for server to shut down.... done
server stopped
```

pghost4□□□□□□□□□□□□□□□□□□

```
...
Oct 16 11:20:35 pghost4 Keepalived_vrrp[7436]:
/usr/local/bin/pg_monitor.sh exited with status 1
Oct 16 11:20:35 pghost4 Keepalived_vrrp[7436]:
VRRP_Script(check_pg_alived) failed
```

```
Oct 16 11:20:35 pghost4 Keepalived_vrrp[7436]:
VRRP_Instance(VI_1) Entering FAULT STATE
Oct 16 11:20:35 pghost4 Keepalived_vrrp[7436]:
VRRP_Instance(VI_1) removing protocol VIPs.
Oct 16 11:20:35 pghost4 Keepalived_vrrp[7436]:
VRRP_Instance(VI_1) Now in FAULT state
Oct 16 11:20:36 pghost4 ntpd[6888]: Deleting interface #6
bond0, 192.168.26.72#123, interface stats: received=0,
sent=0, dropped=0, active_time=442 secs
...
```

□□□□□□check_pg_alived□□□□□Keepalived□□□□□□□VIP□pghost4□bond0□□□□□□□□□□pghost4□□□□□□pghost5□□Keepalived□□□□□□□□□□□□active_standby.sh□□□□□□□□□□□□□□□□□□□□□□□

□□pghost5□□□□□□□□□□

```
...
Oct 16 11:20:38 pghost5 ntpd[6917]: Listen normally on 6
bond0 192.168.26.72 UDP 123
Oct 16 11:20:41 pghost5 Keepalived_vrrp[6561]: Sending
gratuitous ARP on bond0 for 192.168.26.72
Oct 16 11:20:41 pghost5 Keepalived_vrrp[6561]:
VRRP_Instance(VI_1) Sending/queueing gratuitous ARPs on
bond0 for 192.168.26.72
...
```

□□□□□□□pghost5□□Keepalived□□□□□□□□□□□□□□□VIP 192.168.26.72□

pghost5□□□□□/tmp/pg_failover.log□□□□□□

```
2017-10-16 11:20:36: keepalived VIP switchover!
STANDBY_CNT: 1
LAG: 1
Chassis Power Control: Reset
2017-10-16 11:20:37: fence primary db host success.
waiting for server to promote.... done
server promoted
2017-10-16 11:20:37: pghost5 promote standby success.
2017-10-16 11:20:37: pg_failover() function call success.
```

□□□□□□□□Keepalived□VIP□□□□□□□□□□□□□□□□□□□pghost5□□□□□□□□□

□pghost5□□□□□□□□□□□□□□

```
[postgres@pghost5 pg_root]$ pg_controldata | grep cluster
Database cluster state:               in production
```

□□pghost5□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□pghost5□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ipmitool□power reset□□□□pghost5□□□□□□□

□□pghost4□□Keepalived□□□□□□□□□□□ pghost5□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ □□□□□□

□□□□□□□□□□□□□□□□□□□□□□□

·□Keepalived□□□□□□□Keepalived□□□□□□ □□□PostgreSQL□□□□□

·□□□□□□□Keepalived□□□□□□□□□ PostgreSQL□□□□□

·□□□□□□□□□Keepalived□□□□□□□□□ PostgreSQL□□□□□

□□□□□□□□□Keepalived□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□pgpool+□□□□□□□□□ pgpool□□□□□□□□□□□□□□□□□□□□□□□

14.3 [illegible]

[illegible]pgpool[illegible]Keepalived[illegible]pgpool[illegible]pgpool[illegible]pgpool[illegible]Keepalived[illegible]

PostgreSQL[illegible]

[illegible]15[illegible] [illegible]

PostgreSQL[illegible]BUG[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]

15.1 □□□□

□PostgreSQL 10□□PostgreSQL 9.6.14□□□9.6□□□□□14□□□□□□□□□□□□2017□10□□□□PostgreSQL 10□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□10.0□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□2□□5□□8□□11□□□□□□□□□□□□BUG□□□□□□□□□□□□□□□□□□□□

PostgreSQL□□□□□□□□□□□□□□□□Long-term support□LTS□□□□□□□□□□□□□5□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□BUG□□□

PostgreSQL□□□□□□□□□□□□□□□

版　本	发行时间	特　　性
8.4	2009-07-01	窗口函数，列级权限，并行数据库恢复，公用表表达式和递归查询
9.0	2010-09-20	内置流复制、热备、支持 64 位 Windows 操作系统
9.1	2011-09-12	同步流复制，UNLOGGED 表，可串行化快照隔离，可写公用表表达式，SELinux 集成，外部扩展，外部表
9.2	2012-09-10	级联流复制，只用索引的扫描，原生 json 支持，范围类型，pg_receivexlog 工具，Space-Partitioned GiST 索引
9.3	2013-09-09	自定义后台工作进程，数据校验，专用 JSON 运算符，LATERAL JOIN，更快的 pg_dump，pg_isready 服务器监控工具，触发器，视图，可写外部表，物化视图
9.4	2014-12-18	JSONB 数据类型，用于更改配置值的 ALTER SYSTEM 语句，后台工作进程动态注册 / 启动 / 停止，逻辑解析 API，Linux 大页支持，pg_prewarm 缓存预热
9.5	2016-01-07	UPSERT, 行级安全 , 数据抽样 , BRIN 索引
9.6	2016-09-29	并行查询，PostgreSQL 外部表增加排序合并操作下推、多个同步复制从库、vacuum 大表速度增加
10	2017-10-05	逻辑复制，原生分区表、并行查询增强

15.2 □□□□□

PostgreSQL□□□9.6.2□9.6.1□□□□□□□□□□□□9.6.3□9.6.2□□

□□□9.6.4□□□9.6.5□□□□□□□□□□□□□□□□

□□□□□□□□□□□□

□□□□□□□□□□□□□□psql□□□□□□□□□□□□□□□□SELECT version□□□□□□□□□

```
[postgres@pghost1 ~]$ /usr/pgsql-9.6/bin/psql -p 1921 -U
pguser mydb
psql (9.6.4)
Type "help" for help.
mydb=> SELECT version();
        version
------------------------------------------
    PostgreSQL 9.6.4 on x86_64-pc-linux-gnu, compiled by gcc
(GCC) 4.4.7 20120313 (Red Hat 4.4.7-18), 64-bit
(1 row)
```

□□□□□□□□□

[illegible]

[illegible]

[illegible]

```
[postgres@pghost1 ~]$ /usr/pgsql-9.6/bin/psql -U pguser -p
1921 mydb
psql (9.6.5, server 9.6.4)
Type "help" for help.
mydb=> SELECT version();
        version
------------------------------------------
    PostgreSQL 9.6.5 on x86_64-pc-linux-gnu, compiled by gcc
(GCC) 4.4.7 20120313 (Red Hat 4.4.7-18), 64-bit
(1 row)
```

[illegible]psql[illegible]9.6.5[illegible]server[illegible]9.6.4[illegible]psql[illegible]server[illegible]server[illegible]9.6.5[illegible]

15.3 [illegible]

PostgreSQL[illegible]
pg_upgrade[illegible]pg_logical[illegible]PostgreSQL 10[illegible]

[illegible]PostgreSQL[illegible]
release notes[illegible]

15.3.1 [illegible]pg_dumpall[illegible]

[illegible]pg_dumpall[illegible]

[illegible]“[illegible]”[illegible]

[illegible]PostgreSQL 9.3[illegible]10[illegible]

1.[illegible]

□□□□□□□□□□□□□□□RPM□□□□□□□□□□□□□□--prefix□□□□□□□□□□□□□□□PostgreSQL 10□□□□□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/initdb -D /pgdata/10/data/
```

2.□□□□□□□□□□

□□pg_dumpall□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□pg_hba.conf□pg_ident.conf□postgresql.conf□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□pg_hba.conf□pg_ident.conf□□□□□□□□

```
[postgres@pghost1 ~]$ cp /pgdata/9.3/data/pg_hba.conf
/pgdata/10/data/
[postgres@pghost1 ~]$ cp /pgdata/9.3/data/pg_ident.conf
/pgdata/10/data
```

□□□□□□□□□□GUC□□□□□□postgresql.conf□□□□□□□□□□□□□□□□postgresql.conf□□□□□□□□□□□□□□□□□□□□□□□□postgresql.conf□include□□□□□□□□□□□□□□□pg93.conf□□□□□□□□□□□□□□□□□□□□□□□□□□□pg93.conf□□□□□□□□□□□□□□□pg93.conf□□postgresql.conf□□□□□□□□□□

```
[postgres@pghost1 ~]$ cp /pgdata/9.3/data/postgresql.conf
/pgdata/10/data/pg93.conf
```

3.[illegible]

[illegible]

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_ctl -D /pgdata/10/data/
start
waiting for server to start....
2017-10-06 06:51:12.950 GMT [652] LOG:  unrecognized configuration
parameter "checkpoint_segments" in file
"/pgdata/10/data/postgresql.auto.conf" line 28
2017-10-06 06:51:12.950 GMT [652] FATAL:  configuration file
"/pgdata/10/data/postgresql.auto.conf" contains errors
    stopped waiting
pg_ctl: could not start server
```

[illegible]9.3[illegible]10[illegible] release notes[illegible]

4.[illegible]

[illegible] pg_dumpall[illegible]iptables[illegible]pg_hba.conf[illegible]

[illegible]postgresql.conf[illegible] default_transaction_read_only[illegible]on[illegible]

9.4[illegible]ALTER DATABASE[illegible]

```
mydb=# ALTER DATABASE mydb SET default_transaction_read_only = TRUE;
```

[illegible]pgbouncer[illegible]pgbouncer[illegible]

```
[postgres@pghost1 ~]$ /usr/pgsql-9.3/bin/psql -p 5433 -U pguser mydb
psql (9.3.15, server 9.3.15)
Type "help" for help.
mydb=> SELECT state,COUNT(*) FROM pg_stat_activity WHERE pid <>
pg_backend_pid() GROUP BY state;
    state | count
----------+-------
(0 rows)
```

5.[illegible]

[illegible]pg_dumpall[illegible]pg_dumpall[illegible]

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_dumpall -p 5432 >
backup.sql
```

[illegible]

```
/usr/pgsql-10/bin/psql -p 1921 -f backup.sql
```

[illegible]Linux[illegible]

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_dumpall -p 5432 |
/usr/pgsql-10/bin/psql -p 1921
SET
...
CREATE ROLE
ALTER ROLE
ERROR:  role "postgres" already exists
ALTER ROLE
CREATE DATABASE
REVOKE
GRANT
You are now connected to database "mydb" as user "postgres".
...
SET
CREATE EXTENSION
COMMENT
...
ALTER TABLE
COPY 3
 setval
--------
 3
(1 row)
...
```

[illegible]“ERROR[illegible]role"postgres"already exists”[illegible]postgres[illegible]

15.3.2 [illegible]pg_upgrade[illegible]

[illegible]pg_upgrade[illegible]pg_upgrade[illegible]PostGIS[illegible]

□□□PostgreSQL 9.3□□□10□□□□□□□□□
pg_upgrade□□□□□□□

1.pg_upgrade□□

pg_upgrade□□PostgreSQL□□□□□□□□□□□□□□□

□□□□□□□□pg_upgrade□□□□□□□□□□□□□□□

```
-b, --old-bindir=BINDIR       □□□PostgreSQL□□□□□□□□
-B, --new-bindir=BINDIR       □□□PostgreSQL□□□□□□□□
-c, --check                   □□□□□□□□□□□□□□□□
-d, --old-datadir=DATADIR     □□□□□□□□
-D, --new-datadir=DATADIR     □□□□□□□□
-j, --jobs                    □□□□CPU□□□□□□□□□□□□□□□□□□□□□
                              □□□□□□□□□□CPU□□□□□□□□□□□□□□□□
                              □□□
-k, --link                    □□□□□□□
-o, --old-options=OPTIONS     □□□□□□postgres □□□□□□,□□□□□□□□□□□
-O, --new-options=OPTIONS     □□□□□□postgres □□□□□□,□□□□□□□□□□□
-p, --old-port=PORT           □□□□□□□□□□pg_upgrade□□□□□50432□
                              □□□□□□□□□□□□□□□
-P, --new-port=PORT           □□□□□□□□□□□□□□□□□□□□□□□□□□□□
                              □□□□□□□□□□□50432□□□□□□□□□□□□□□
                              □□□□□□□□□□□,□□□□□□□□□□□□□□□□□
-r, --retain                  □□□□□□□□□□□SQL□□□□
```

□□□□□□□□□pg_upgrade□□-c□□pg_upgrade□□□-c□□

pg_upgrade□□□□□□Link□□

□□□□□□link□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□

2.□□pg_upgrade□□

1□□□□□□PostgreSQL□□□□□□□□□

2□□□□□□□□□□

3□□□□□□□□□□□

□□□□□□□□“--check”□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_upgrade -b /usr/pgsql-
9.3/bin -B /usr/pgsql-10/bin -d /pgdata/9.3/data/ -D /pgdata/10/data/
-k -c
Performing Consistency Checks
-----------------------------
Checking cluster versions                                   ok
...
Checking for prepared transactions                          ok
*Clusters are compatible*
```

□□□□□□“Clusters are compatible”□□□□□□□□□□
□□□□□□□□□□“Failure□exiting”□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□

```
Your installation references loadable libraries that are missing from
the
new installation.  You can add these libraries to the new
installation,
or remove the functions using them from the old installation.  A list
of
problem libraries is in the file:
    loadable_libraries.txt
```

□□□□□□loadable_libraries.txt□□□□□□□□

```
[postgres@pghost1 ~]$ cat loadable_libraries.txt
could not load library "$libdir/postgis-2.3": ERROR:  could not
access file "$libdir/postgis-2.3": No such file or directory
could not load library "$libdir/rtpostgis-2.3": ERROR:  could not
access file "$libdir/rtpostgis-2.3": No such file or directory
```

□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□warning□□□□□□

```
Checking for hash indexes warning
Your installation contains hash indexes.  These indexes have
different
internal formats between your old and new clusters, so they must be
reindexed with the REINDEX command.  After upgrading, you will be
given
REINDEX instructions.
```

□□□□□PostgreSQL 10□□□□□□□□hash□□□□□□□□□□□□□□□□□□□hash□□□□□□□□□□□□

```
mydb=# \d+ tbl
...
Indexes:
    "tbl_hash_idx" hash (column_name) INVALID
```

□□□□warning□□□□□□□□□□□□□□□hash□□□□□□□□□□□□□□□□□□□□

```
mydb=# CREATE INDEX CONCURRENTLY ON th USING HASH(column_name);
CREATE INDEX
```

```
mydb=# DROP INDEX tbl_hash_idx;
DROP INDEX
```

4□□□pg_upgrade□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_upgrade -b /usr/pgsql-
9.3/bin -B /usr/pgsql-10/bin -d /pgdata/9.3/data/ -D /pgdata/10/data/
Performing Consistency Checks
-----------------------------
Checking cluster versions                                   ok
Checking database user is the install user                  ok
Checking database connection settings                       ok
Checking for prepared transactions                          ok
Checking for reg* data types in user tables                 ok
Checking for contrib/isn with bigint-passing mismatch       ok
Checking for invalid "unknown" user columns                 ok
Checking for roles starting with "pg_"                      ok
Checking for incompatible "line" data type                  ok
Creating dump of global objects                             ok
Creating dump of database schemas
                                                            ok
Checking for presence of required libraries                 ok
Checking database user is the install user                  ok
Checking for prepared transactions                          ok
If pg_upgrade fails after this point, you must re-initdb the
new cluster before continuing.
Performing Upgrade
------------------
Analyzing all rows in the new cluster                       ok
Freezing all rows in the new cluster                        ok
Deleting files from new pg_xact                             ok
Copying old pg_clog to new server                           ok
Setting next transaction ID and epoch for new cluster       ok
Deleting files from new pg_multixact/offsets                ok
Copying old pg_multixact/offsets to new server              ok
Deleting files from new pg_multixact/members                ok
Copying old pg_multixact/members to new server              ok
Setting next multixact ID and offset for new cluster        ok
Resetting WAL archives                                      ok
Setting frozenxid and minmxid counters in new cluster       ok
Restoring global objects in the new cluster                 ok
Restoring database schemas in the new cluster
                                                            ok
Copying user relation files
                                                            ok
Setting next OID for new cluster                            ok
Sync data directory to disk                                 ok
Creating script to analyze new cluster                      ok
```

```
Creating script to delete old cluster                       ok
Checking for hash indexes                                   ok
Upgrade Complete
----------------
Optimizer statistics are not transferred by pg_upgrade so,
once you start the new server, consider running:
    ./analyze_new_cluster.sh
Running this script will delete the old cluster's data files:
    ./delete_old_cluster.sh
```

□□□□pg_upgrade□□□□□□□□□□□□□□□□□□□□□“Upgrade Complete”□□□□□□□□□□□

5□□□pg_upgrade□link□□□□□

□□link□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Extension□□□□□□□□□□

```
mydb=# \db
            List of tablespaces
        Name     |  Owner   |       Location
-----------------+----------+---------------------
    pg_default   | postgres |
    pg_global    | postgres |
    pgtablespace | postgres | /pgdata/pgtablespace
(3 rows)
mydb=# \dx
                                     List of installed extensions
       Name        | Version |   Schema   |
Description
-------------------+---------+------------+-------------------------
-------
pg_stat_statements | 1.1     | public     | track execution
statistics of all SQL statements executed
plpgsql            | 1.0     | pg_catalog | PL/pgSQL procedural
language
(2 rows)
```

□□pg_upgrade□□□□□□□□

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_upgrade -b /usr/pgsql-
9.3/bin -B /usr/pgsql-10/bin -d /pgdata/9.3/data/ -D /pgdata/10/data/
-k
Performing Consistency Checks
-----------------------------
Checking cluster versions                                   ok
...
Checking for prepared transactions                          ok
If pg_upgrade fails after this point, you must re-initdb the
new cluster before continuing.
Performing Upgrade
------------------
Analyzing all rows in the new cluster                       ok
Freezing all rows in the new cluster                        ok
Deleting files from new pg_xact                             ok
Copying old pg_clog to new server                           ok
Setting next transaction ID and epoch for new cluster       ok
Deleting files from new pg_multixact/offsets                ok
Copying old pg_multixact/offsets to new server              ok
Deleting files from new pg_multixact/members                ok
Copying old pg_multixact/members to new server              ok
Setting next multixact ID and offset for new cluster        ok
Resetting WAL archives                                      ok
Setting frozenxid and minmxid counters in new cluster       ok
Restoring global objects in the new cluster                 ok
Restoring database schemas in the new cluster
                                                            ok
Adding ".old" suffix to old global/pg_control               ok
If you want to start the old cluster, you will need to remove
the ".old" suffix from /pgdata/9.3/data/global/pg_control.old.
Because "link" mode was used, the old cluster cannot be safely
started once the new cluster has been started.
Linking user relation files
                                                            ok
Setting next OID for new cluster                            ok
Sync data directory to disk                                 ok
Creating script to analyze new cluster                      ok
Creating script to delete old cluster                       ok
Checking for hash indexes                                   ok
Upgrade Complete
----------------
Optimizer statistics are not transferred by pg_upgrade so,
once you start the new server, consider running:
    ./analyze_new_cluster.sh
Running this script will delete the old cluster's data files:
    ./delete_old_cluster.sh
```

[illegible]pg_upgrade[illegible]pg_upgrade[illegible]pg_control[illegible]pg_control.old[illegible]pg_control.old[illegible]pg_control[illegible]

6[illegible]

pg_upgrade[illegible]

[illegible]pg_upgrade[illegible]analyze_new_cluster.sh[illegible]vacuumdb--analyze-in-stages[illegible]VACUUM[illegible]VACUUM[illegible]

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/vacuumdb -a --analyze-in-
stages -h pghost1 -p 1921
```

7[illegible]

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_ctl -D /pgdata/10/data/
start

[postgres@pghost1 ~]$ /usr/pgsql-10/bin/psql -p 1921 mydb
psql (10.0)
Type "help" for help.
mydb=# SELECT version();
        version
-----------------------------------------
PostgreSQL 10.0 on x86_64-pc-linux-gnu, compiled by gcc (GCC) 4.4.7
20120313 (Red Hat 4.4.7-18), 64-bit
(1 row)
```

8[illegible]

[illegible]

3.[illegible]pg_upgrade[illegible]

[illegible]

1[illegible]

PostgreSQL[illegible]

```
postgres=# SELECT application_name,client_addr,sync_state FROM
pg_stat_replication;
    application_name | client_addr  | sync_state
---------------------+--------------+------------
    walreceiver      | 10.191.136.3 | async
(1 row)
```

[illegible]sync_state[illegible]async[illegible]sync[illegible]

[illegible]postgresql.conf[illegible]“synchronous_standby_names”[illegible]*[illegible]

[illegible]reload[illegible]

```
mydb=# SELECT pg_reload_conf();
 pg_reload_conf
----------------
 t
(1 row)
```

2□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□“Latest checkpoint location”□□□□□□□□

```
[postgres@pghost1 ~]$ /usr/pgsql-9.3/bin/pg_ctl -D
/pgdata/9.3/data_master -m fast stop
waiting for server to shut down.... done
server stopped
[postgres@pghost1 ~]$
```

□□pg_controldata□□□□□□□□“Latest checkpoint location”□□□□□□□□□□

```
[postgres@pghost1 ~]$ /usr/pgsql-9.3/bin/pg_controldata
/pgdata/9.3/data_master |
grep "Latest checkpoint location"
Latest checkpoint location:           0/1F000028
[postgres@pghost2 ~]$ /usr/pgsql-9.3/bin/pg_controldata
/pgdata/9.3/data_slave | grep "Latest checkpoint location"
Latest checkpoint location:           0/1F000028
```

3□□□□□□□□□□

□□□□□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/initdb -k -D
/pgdata/10/data_master
```

[illegible]“Database system identifier”[illegible]

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_controldata
/pgdata/10/data_master | grep "Database system identifier"
Database system identifier:           6504563477800205659
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/pg_controldata
/pgdata/10/data_slave | grep "Database system identifier"
Database system identifier:           6504563477800205659
```

[illegible]pg_upgrade[illegible]

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_upgrade -b /usr/pgsql-
9.3/bin -B /usr/pgsql-10/bin -d /pgdata/9.3/data_master/ -D
/pgdata/10/data_master/ -k
```

[illegible]

4[illegible]

[illegible]

```
[postgres@pghost2 ~]$ cp /pgdata/9.3/data_slave/*.conf
/pgdata/9.3/backup/
```

[illegible]pg_upgrade[illegible]

```
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/pg_upgrade -b /usr/pgsql-
9.3/bin -B /usr/pgsql-10/bin -d /pgdata/9.3/data_slave/ -D
/pgdata/10/data_slave/ -k
```

[illegible]“Database system identifier”[illegible]

```
2017-11-28 18:42:24.692 CST,,,82112,,5a44ca90.140c0,1,,2017-11-28
18:42:24 CST,,0,FATAL,XX000,"database system identifier differs
between the primary and standby","The primary's identifier is
6504542580012252832,
the standby's identifier is
6504544106935856722.",,,,,,,"WalReceiverMain, walreceiver.c:347",""
```

5[illegible]

[illegible]Master[illegible]

```
[postgres@pghost1 ~]$ /usr/pgsql-10/bin/pg_ctl -D
/pgdata/10/data_master start
```

6[illegible]

[illegible]recovery.conf[illegible]

```
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/pg_ctl -D
/pgdata/10/data_slave start
```

7[illegible]

[illegible]

```
mydb=# select version();
version
```

```
----------------------------------------
    PostgreSQL 10.1 on x86_64-pc-linux-gnu, compiled by gcc (GCC)
4.4.7 20120313 (Red Hat 4.4.7-18), 64-bit
(1 row)
```

□□□□□□□□□□□□□□□□□□

```
mydb=# SELECT client_addr,sync_state FROM pg_stat_replication;
    client_addr | sync_state
----------------+------------
    pghost2     | sync
(1 row)
```

15.3.3 □□pglogical□□□□□

1.pglogical□□

2ndQuadrant□□□□□pglogical Extension□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□pglogical□□□□Node□□□□PostgreSQL□□□Providers□Subscribers□□□Node□□□□Replication Set□□□□□□□□□□□

2.□□pglogical□□□□□

□□pglogical□□

pglogical□□□□□□

·PostgreSQL 9.4□□□□□□

·□□□□□□□PRIMARY KEY□□REPLICA IDENTITY□

·□□□□UNLOGGED□TEMPORARY□□

·□□□□□□□□□□□□□Database□

·□□□□DDL□□□□□□□DDL□□□
pglogical.replicate_ddl_command□□□□□□□□□□□□□

□□□□□□□□□□□□□□pglogical□□pglogical□□□□□

3.□□pglogical□□□□□□□

□□□PostgreSQL 9.4□□□PostgreSQL 10□□□□□□□□□pglogical□□□□□□

1□□□□□□□□□□□□□□□□□

```
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/initdb -k -D
/pgdata/10/data_master
```

2□□□□□□□□□□□□

```
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/pg-ctl -D
/pgdata/10/data_master start
```

3□□□□□□□schema□□□□□□□□□□□

```
[postgres@pghost2 ~]$ /usr/pgsql-10/bin/pg_dump -Fp -v -c -s -h
pghost1 -p 1921 -d mydb |
/usr/pgsql-10/bin/psql -h pghost2 -p 1922 -d mydb
```

4□□□pglogical□

□□□□□□□□□pglogical□repo RPM□□□□□□

```
[root@pghost1 ~]# yum install
http://packages.2ndquadrant.com/pglogical/yum-repo-rpms/pglogical-
rhel-1.0-3.noarch.rpm
```

□□□□□□□□□□pglogical□□□□□□

```
# PostgreSQL 9.4
[root@pghost1 ~]# yum install postgresql94-pglogical.x86_64
# PostgreSQL 10
[root@pghost2 ~]# yum install postgresql10-pglogical.x86_64
```

5□□□pglogical□□□□□

□PostgreSQL 9.4□□□□□□□□□□

·□□pg_hba.conf□□□□□

·□□postgresql.conf□□□wal_max_sender□□□□□□0□□□10□

·□□postgresql.conf□□□wal_level□□□logical□

·□□postgresql.conf□□□max_replication_slots□□□□□□0□□□10□

□□□□□□□postgresql.conf□ shared_preload_libraries□□□□□□pglogical□□□ shared_preload_libraries□□□□□□□□□□□□□□□□□□□□

```
shared_preload_libraries = 'pglogical'
```

6□□□pglogical Extension□

□□□□□□□□□pglogical Extension□□□□□ PostgreSQL 9.4□□□□□□□pglogical_origin Extension□□□□□□

```
mydb=# SELECT version();
        version
------------------------------------------------
    PostgreSQL 9.4.15 on x86_64-unknown-linux-gnu, compiled by gcc
(GCC) 4.8.5 20150623 (Red Hat 4.8.5-16), 64-bit
(1 row)
mydb=# CREATE EXTENSION pglogical_origin;
CREATE EXTENSION
mydb=# CREATE EXTENSION pglogical;
CREATE EXTENSION
mydb=# \dx pglogical
        List of installed extensions
    Name      | Version |  Schema   |          Description
--------------+---------+-----------+-------------------------------
    pglogical | 2.1.0   | pglogical | PostgreSQL Logical Replication
(1 row)
```

□□□□□□pglogical Extension□□□□□□

```
mydb=# SELECT version();
    version
-----------------------------------------
    PostgreSQL 10.1 on x86_64-pc-linux-gnu, compiled by gcc (GCC)
4.8.5 20150623 (Red Hat 4.8.5-16), 64-bit
(1 row)
```

```
mydb=# CREATE EXTENSION pglogical;
CREATE EXTENSION
mydb=# \dx pglogical
                   List of installed extensions
      Name    | Version |  Schema   |          Description
--------------+---------+-----------+--------------------------------
    pglogical | 2.1.0   | pglogical | PostgreSQL Logical Replication
(1 row)
```

7□□□□□□□provider node□□□□□□

```
mydb=# SELECT pglogical.create_node(node_name := 'pg94provider', dsn
:= 'host=127.0.0.1 port=1921 dbname=mydb');
    create_node
-------------
    3412564209
(1 row)
```

□□pglogical□□□pglogical_node_info□□□□□□□□provider node□□□□□□□□□

```
mydb=# SELECT * FROM pglogical.pglogical_node_info();
node_id   |  node_name    |        sysid         | dbname |
replication_sets
----------+--------------+---------------------+--------+-----------
---------
3412564209| pg94provider | 6488108006998879651 | mydb   |
".\x0Bg0\x02",",vg\x18\x02","Lg\x02"
(1 row)
```

8□□provider node□□□□□□□□□□□□□□

```
mydb=# SELECT
pglogical.create_replication_set('insert_update_delete_notruncate',TR
UE,TRUE,TRUE,FALSE);
    create_replication_set
------------------------
        798613796
(1 row)
```

```
mydb=# SELECT * FROM pglogical.replication_set WHERE set_name =
'insert_update_delete_notruncate';
    set_id   | set_nodeid |            set_name             |
replicate_insert | replicate_update | replicate_delete |
replicate_truncate
    -----------+------------+---------------------------------+------
------------+------------------+------------------+------------------
--
    798613796 | 3412564209 | insert_update_delete_notruncate | t
| t                | t                | f
(1 row)
```

9□□provider node□□□□□□□□□□□□

```
mydb=# SELECT
pglogical.replication_set_add_all_tables('insert_update_delete_notrun
cate', ARRAY['public']);
    replication_set_add_all_tables
--------------------------------
    t
(1 row)
```

10□□PostgreSQL 10□□□□subscription node□□□□□□

```
mydb=# SELECT pglogical.create_node(node_name := 'pg10subscriber',dsn
:= 'host= 127.0.0.1 port=1922 dbname=mydb');
    create_node
-------------
    3105221948
(1 row)
mydb=# SELECT pglogical.create_subscription(subscription_name :=
'pg10subscription', provider_dsn := 'host=127.0.0.1 port=1921
dbname=mydb');
    create_subscription
---------------------
        1639099831
(1 row)
```

11□□subscription node□□□□□□□□□□

```
mydb=# SELECT
pglogical.alter_subscription_synchronize('pg10subscription',FALSE);
 alter_subscription_synchronize
--------------------------------
 t
(1 row)
```

□□□□□□“finished sync of table xxx for subscriber pg10subscription”□□□□□□□□□□□□□□□□□□□

```
mydb=# SELECT pglogical.show_subscription_status('pg10subscription');
show_subscription_status
---------------------------------
    (pg10subscription,replicating,pg94provider,"host=127.0.0.1
port=1921 dbname=mydb",pgl_mydb_pg94provider_pg10subscription,
{insert_update_delete_notruncate},{all})
   (1 row)
```

□□□□□□□□□□□□□□□□□□□□□□

12□□□□□□□□□□□□□□□□□

```
mydb=# ALTER SYSTEM SET default_transaction_read_only = 'on';
ALTER SYSTEM
```

13□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□github□□□□□□□pglogical□□□□□□□□□□□□□□□□□□□□□□□□□□□□

15.4 [illegible]

[illegible]PostgreSQL[illegible]pg_dumpall[illegible]pglogical[illegible]pg_upgrade[illegible]pg_upgrade[illegible]pg_upgrade[illegible]PostgreSQL 10[illegible]PostgreSQL 10[illegible]

[illegible]

□16□ □□□□

PostgreSQL□□□□□□□□□□□□□□□□□□
PostgreSQL□□□□□□□□□□□□□□□□□

·□□□□PostgreSQL□□□world□□□□□□□□□

·□□GitHub□□□□□□□□□□□□

□□□□□□□□□□50□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□PostgreSQL 9.0□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□

16.1 CREATE EXTENSION

[illegible]gmake[illegible]PostgreSQL[illegible]world[illegible]$PGHOME/share/extension/[illegible]PostgreSQL[illegible]

```
# gmake world
# gmake install-world
```

[illegible]$PGHOME/share/extension/[illegible]

```
$ ll $PGHOME/share/extension/
...
-rw-r--r-- 1 root root  794 Oct  6 10:20 pg_buffercache-
-1.2.sql
-rw-r--r-- 1 root root  157 Oct  6 10:20
pg_buffercache.control
...
```

[illegible].control[illegible]gmake world[illegible]

[illegible]gmake world[illegible]PostgreSQL[illegible]CREATE EXTENSION[illegible]

```
CREATE EXTENSION [ IF NOT EXISTS ] extension_name
```

[illegible]PostgreSQL 9.1[illegible]CREATE EXTENSION[illegible]SQL[illegible]

[illegible]pg_buffercache[illegible]

```
postgres=# CREATE EXTENSION pg_buffercache;
CREATE EXTENSION
```

[illegible]\dx[illegible]

```
postgres=# \dx
            List of installed extensions
        Name       | Version |   Schema   |
Description
-------------------+---------+------------+----------------
----------------
    pg_buffercache | 1.3     | public     | examine the
shared buffer cache
    plpgsql        | 1.0     | pg_catalog | PL/pgSQL
procedural language
(2 rows)
```

[illegible]pg_extension[illegible]

```
postgres=# SELECT * FROM pg_extension ;
extname     | extowner | extnamespace | extrelocatable |
extversion | extconfig | extcondition
-------------+---------+-----------+-------------+------
-----+----------+
plpgsql        |       10 |           11 | f              |
1.0       |           |
pg_buffercache |       10 |         2200 | t              |
1.3       |           |
(2 rows)
```

[illegible]pg_available_extensions[illegible]

```
pg_available_extensions[illegible]
postgres=# SELECT * FROM pg_available_extensions;
    name        | default_version | installed_version |
comment
----------------+-----------------+------------------+------
---------------
...
pg_buffercache    | 1.3               | 1.3               |
examine the shared buffer cache
...
(43 rows)
```

pg_available_extensions[illegible]

·name[illegible]

·default_version[illegible]

·installed_version[illegible]

·comment[illegible]

[illegible]

16.2 pg_stat_statements

pg_stat_statements□□□□□□□□□□□□□SQL□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□□□SQL□□□□□□□□□□□□□□□□□□□

pg_stat_statements□□□□□□□□□□□□□□□□□postgresql.conf□□□□□□□shared_preload_libraries□□□□pg_stat_statements□□□□□□

```
shared_preload_libraries = 'pg_stat_statements'         # (change requires restart)
pg_stat_statements.max = 10000
pg_stat_statements.track = all
pg_stat_statements.track_utility = on
pg_stat_statements.save = on
```

shared_preload_libraries□□□pg_stat_statements□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□pg_stat_statements.□□□□□□□pg_stat_statements□□□□□postgresql.conf□□□□□□□□□□□□□□□□□□□□□

·pg_stat_statements.max□pg_stat_statements□□□□□□□□□□□□□SQL□□□□5000□□□□□□□□□□□□□□□□□SQL□□□□□

·pg_stat_statements.track□pg_stat_statements□□□□□□□SQL□□□□top□□□□□SQL□all□□□□□□□□SQL□□□□□□all□

·pg_stat_statements.track_utility□pg_stat_statements□□□□□□□□□SELECT□UPDATE□DELETE□INSERT□□□SQL□□□□□□on□

·pg_stat_statements.save□pg_stat_statements□□□□□□□□□□□□□□□SQL□□□□□□□□□off□□□□□□□□□SQL□□□□□□□□□□□□□□on□

□□□□□□□□□□□□□

```
[postgres@pghost1 pg_root]$ pg_ctl restart -m fast
```

□postgres□□□□□□□□□□□□□□□□postgres□□□□□□pg_stat_statements□□□□□□□□

```
[postgres@pghost1 loadtest]$ psql postgres
psql (10.0)
Type "help" for help.
```

```
postgres=# CREATE EXTENSION pg_stat_statements ;
CREATE EXTENSION
```

□□□□□pg_stat_statements□□□pg_stat_statements_reset□□□□□□□□□□□

pg_stat_statements□□□□□□□□□□16-1□□□

□16-1 pg_stat_statements□□□□□□

名　称	类　型	关联信息	描　述
userid	oid	pg_authid.oid	执行此 SQL 的用户 OID
dbid	oid	pg_database.oid	此 SQL 执行的数据库 OID
queryid	bigint		此 SQL 的编号
query	text		此 SQL 内容
calls	bigint		此 SQL 调用次数
total_time	double precision		此 SQL 执行的总时间，单位毫秒
min_time	double precision		此 SQL 执行的最小时间，单位毫秒
max_time	double precision		此 SQL 执行的最大时间，单位毫秒
mean_time	double precision		此 SQL 执行的平均时间，单位毫秒
rows	bigint		此 SQL 影响的数据行
shared_blks_hit	bigint		此 SQL 命中的共享内存数据块数
shared_blks_read	bigint		此 SQL 读取的共享内存数据块数
shared_blks_dirtied	bigint		此 SQL 产生的共享内存数据脏块数量
shared_blks_written	bigint		此 SQL 写入的共享内存数据块数

□□□□□□pg_stat_statements□□□□□□□□□□□□□□□□□□□□□□□□□pgbench□□□□SQL□□□□□□□□□SQL□□□□□□□□pg_stat_statements□□□□

[illegible]tran_per1.sql[illegible]

```
\set v_id random(1,10000000)

SELECT name FROM test_per1 WHERE id=:v_id;
UPDATE test_per2 SET flag='1' WHERE id=:v_id;
```

[illegible]pgbench[illegible]SQL[illegible]

```
$ pgbench -c 4 -T 120 -d postgres -U postgres -n N -M
prepared -f tran_per1.sql >
tran_per1_4.out 2>&1 &
```

[illegible]pg_stat_statements[illegible]SQL[illegible]SQL[illegible]SQL[illegible]calls[illegible]

```
postgres=# SELECT userid,dbid,queryid,query,calls,
    total_time,min_time,max_time,mean_time,rows
 FROM pg_stat_statements ORDER BY calls DESC LIMIT 2;
-[ RECORD 1 ]---------------------------------------
userid     | 10
dbid       | 13158
queryid    | 1758110241
query      | UPDATE test_per2 SET flag=$2 WHERE id=$1
calls      | 812691
total_time | 83537.8480209997
min_time   | 0.015113
max_time   | 6583.11359
mean_time  | 0.102791649004361
rows       | 812691
-[ RECORD 2 ]---------------------------------------
userid     | 10
dbid       | 13158
```

```
queryid    | 2354798721
query      | SELECT name FROM test_per1 WHERE id=$1
calls      | 812691
total_time | 28687.6311289994
min_time   | 0.008581
max_time   | 76.785687
mean_time  | 0.0352995555863171
rows       | 812691
```

□□□SQL□□□□□□

```
postgres=# SELECT userid,dbid,queryid,query,calls,

total_time,min_time,max_time,mean_time,rows
FROM pg_stat_statements ORDER BY mean_time DESC LIMIT 1;
-[ RECORD1 ]------------------------------------------
userid     | 10
dbid       | 13158
queryid    | 125206108
query      | select pg_sleep($1)
calls      | 1
total_time | 6019.78287
min_time   | 6019.78287
max_time   | 6019.78287
mean_time  | 6019.78287
```

□□□SQL□□□□mean_time□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□

□□□□□□□□pg_stat_statements□□□□□□□□□□□□pg_stat_statements_reset□□□□□□□□□□SQL□□□□□□□□

```
postgres=# SELECT COUNT(*) FROM pg_stat_statements ;
    count
-------
        20
(1 row)

postgres=# SELECT pg_stat_statements_reset();
    pg_stat_statements_reset
--------------------------

(1 row)

postgres=# SELECT COUNT(*) FROM pg_stat_statements ;
 count
-------
 2
(1 row)
```

---

pg_stat_statements[illegible]CREATE EXTENSION[illegible]

16.3 auto_explain

□□□□□□□□EXPLAIN□EXPLAIN ANALYZE□□□□SQL□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□auto_explain□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□SQL□□□□□□□□□□□

auto_explain□□□□□pg_stat_statements□□□□□□□□□□□□postgresql.conf□□□□□□□shared_preload_libraries□□□□□□□□

```
shared_preload_libraries = 'pg_stat_statements,auto_explain'
auto_explain.log_min_duration = 0
auto_explain.log_analyze = on
auto_explain.log_buffers = off
```

□auto_explain□□□□□□auto_explain□□□□□postgresql.conf□□□□□□□□□□□□□□□□□□□□□

·auto_explain.log_min_duration□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□0□□□□□□SQL□□□□□□□□□-1□□□□□□□

[illegible]100ms[illegible]100[illegible]SQL[illegible]

·auto_explain.log_analyze[illegible]SQL[illegible]ANALYZE[illegible]EXPLAIN[illegible]ANALYZE[illegible]off[illegible]

·auto_explain.log_buffers[illegible]SQL[illegible]EXPLAIN[illegible]BUFFERS[illegible]off[illegible]

[illegible]auto_explain[illegible]EXPLAIN[illegible]EXPLAIN[illegible]

[illegible]

```
[postgres@pghost1 pg_log]$ pg_ctl restart -m fast
```

[illegible]

[illegible]SQL[illegible]SQL[illegible]

```
postgres=# SELECT * FROM test_per1 WHERE id=1;
    id |  name  |     create_time
-------+--------+--------------------
     1 | 1_per1 | 2017-09-04 21:21:44
(1 row)
```

□□□□□□□□□□□□□□□

```
    [postgres@pghost1 pg_log]$ tail -f postgresql-2017-10-
22_201048.csv
2017-10-22 20:18:25.927 CST,"postgres","postgres",24856,"
[local]",59ec8c8d.6118,1,"SELECT",2017-10-22 20:18:21
CST,5/367,0,LOG,00000,"duration: 0.049 ms  plan:
Query Text: SELECT * FROM test_per1 WHERE id=1;
Index Scan using test_per1_pkey on test_per1  (cost=0.43..4.45
rows=1 width=24) (actual time=0.023..0.023 rows=1 loops=1)
    Index Cond: (id = 1)",,,,,,,,,"psql"
```

□□□□□□□□□□□□□SQL□□□□SQL□□□□□□□□

□□□□□□□□□□□□auto_explain.log_min_duration□□□□0□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□100□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□

auto_explain□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□SQL□□□

16.4 pg_prewarm

[illegible]PostgreSQL9.4[illegible]pg_prewarm[illegible]

[illegible]postgres[illegible]mydb[illegible]pg_prewarm[illegible]

```
[postgres@pghost1 ~]$ psql mydb postgres
psql (10.0)
Type "help" for help.

mydb=# CREATE EXTENSION pg_prewarm;
CREATE EXTENSION
```

[illegible]pg_prewarm[illegible]pg_prewarm[illegible]

```
pg_prewarm(regclass, mode text default 'buffer', fork text
default 'main',
    first_block int8 default null,
    last_block int8 default null) RETURNS int8
```

[illegible]

·regclass[illegible]

·mode[illegible]prefetch[illegible]read[illegible]buffer[illegible]

·fork[illegible]main[illegible]

·first_block[illegible]null[illegible]0[illegible]

·last_block[illegible]null[illegible]

[illegible]mydb[illegible]t_pre[illegible]200[illegible]

```
[postgres@pghost1 ~]$ psql mydb pguser
mydb=> CREATE TABLE t_pre(id int4 PRIMARY KEY,
        info text,
        create_time timestamp(0) without time zone);
CREATE TABLE

mydb=> INSERT INTO t_pre(id,info,create_time)
    SELECT n,n||'_pre',clock_timestamp() FROM
generate_series(1,2000000) n ;
INSERT 0 2000000
```

[illegible]t_pre[illegible]

```
mydb=> SELECT pg_prewarm('t_pre','buffer');
 pg_prewarm
------------
 12739
(1 row)
```

[illegible]12739[illegible]t_pre[illegible]pg_class[illegible]

```
mydb=> SELECT relname,relpages FROM pg_class WHERE
relname='t_pre';
    relname | relpages
------------+----------
    t_pre   |    12739
(1 row)
```

relpages[illegible]ANALYZE[illegible]

[illegible]

```
mydb=> SELECT pg_prewarm('t_pre','prefetch');
 pg_prewarm
------------
 12739
(1 row)
```

[illegible]

```
mydb=> SELECT pg_prewarm('t_pre','read');
    pg_prewarm
```

```
------------
        12739
(1 row)
```

[illegible]pg_prewarm[illegible]pg_prewarm[illegible]

[illegible]pgfincore[illegible]PostgreSQL9.4[illegible]pg_prewarm[illegible]pgfincore[illegible]https://github.com/klando/pgfincore[illegible]

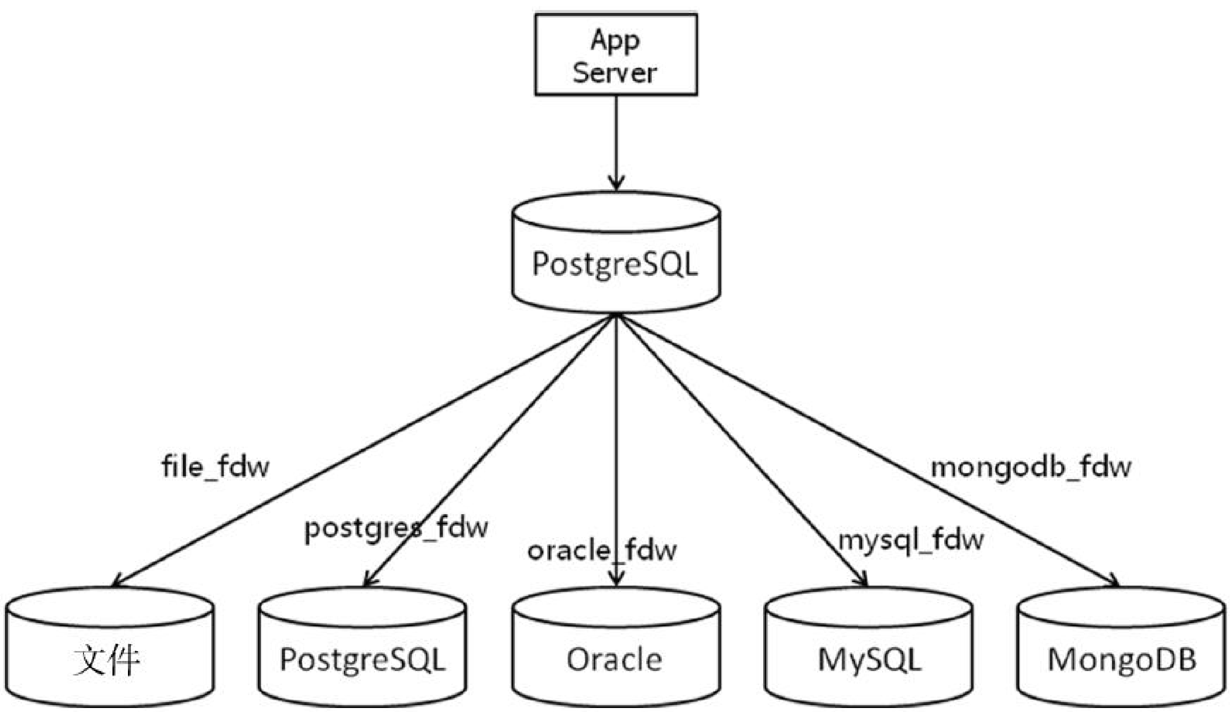
[illegible] [illegible]pg_prewarm[illegible]pg_prewarm[illegible]

16.5 file_fdw

file_fdw□□□PostgreSQL□□□□□□□□file_fdw□□□□□□□SQL/MED□SQL Management of External Data□□

16.5.1 SQL/MED□□

SQL/MED□PostgreSQL□□□□□□□□□□□□□□PostgreSQL□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□Oracle□□dblink□□□□□SQL/MED□□□□16-1□□□



[illegible]16-1 SQL/MED[illegible]

[illegible]

·[illegible]PostgreSQL[illegible]csv[illegible]text[illegible]

·[illegible]PostgreSQL[illegible]PostgreSQL[illegible]Oracle[illegible]MySQL[illegible]SQL Server[illegible]

·[illegible]MongoDB[illegible]Redis[illegible]Cassandra[illegible]

·[illegible]Elastic Search[illegible]Hadoop[illegible]

[illegible]PostgreSQL[illegible]

16.5.2 file_fdw[illegible]

[illegible]PostgreSQL[illegible]file_fdw[illegible]text[illegible]csv[illegible]binary[illegible]

[illegible]file_fdw[illegible]

1[illegible]file_fdw[illegible]

2[illegible]foreign server[illegible]

3[illegible]file_fdw[illegible]

4[illegible]

[illegible]postgres[illegible]mydb[illegible]file_fdw[illegible]

```
[postgres@pghost1 ~]$ psql mydb postgres
psql (10.0)
Type "help" for help.

mydb=# CREATE EXTENSION file_fdw;
CREATE EXTENSION
```

[illegible]file_fdw[illegible]fs_file[illegible]

```
mydb=# CREATE SERVER fs_file FOREIGN DATA WRAPPER file_fdw;
CREATE SERVER
```

[illegible]fs_file[illegible]\des[illegible]

```
mydb=# \des
            List of foreign servers
        Name|  Owner   | Foreign-data wrapper
------------+----------+---------------------
    fs_file | postgres | file_fdw
(1 row)
```

□□IP□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□/home/postgres/script/file1.txt□□□□□□□□□□tab□□□□

```
1       a
2       b
```

□□□□file_fdw□□□□□□□□□□

```
mydb=# CREATE FOREIGN TABLE ft_file1(
        id int4,
        flag text
    ) SERVER fs_file
        OPTIONS (filename
'/home/postgres/script/file1.txt',format 'text');
CREATE FOREIGN TABLE
```

OPTIONS□□file_fdw□□□□□□□□□fdw□□□□□□□□file_fdw□□□□□□□□

·filename[illegible]

·format[illegible]text[illegible]csv[illegible]binary[illegible]text[illegible]

·header[illegible]csv[illegible]

·delimiter[illegible]text[illegible]tab[illegible]

·encoding[illegible]

[illegible]filename[illegible]copy[illegible]

[illegible]

```
mydb=# SELECT * FROM ft_file1 ;
    id | flag
-------+------
     1 | a
     2 | b
(2 rows)
```

[illegible]file_fdw[illegible]

[illegible]PostgreSQL[illegible]

[illegible]file_fdw[illegible]INSERT/UPDATE/DELETE[illegible]

[illegible]\det[illegible]

```
mydb=# \det
        List of foreign tables
    Schema |  Table   | Server
-----------+----------+---------
    public | ft_file1 | fs_file
(1 row)
```

16.5.3 [illegible]file_fdw[illegible]

[illegible]file_fdw[illegible]PostgreSQL[illegible]csv[illegible]csv[illegible]postgresql.conf[illegible]

```
log_destination = 'csvlog'
logging_collector = on
log_directory = 'pg_log'
log_filename = 'postgresql-%Y-%m-%d_%H%M%S.log'
```

[illegible]

·log_destination□□□postgresql□□□□□□□□□□□□stderr□csvlog□syslog□□□□□□□□csv□□□csv□□□□□logging_collector□□□on□□

·logging_collector□□□□□□□on□□□□□□□□□□□□□□□□□□stderr□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

·log_directory□□logging_collector□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□$PGDATA□□

·log_filename□□logging_collector□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□

$PGDATA/pg_log/□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□postgresql-2017-10-26_000000.csv□□□□□□□□□□□□□□□

```
...
2017-10-26 14:29:16.298 CST,"postgres","mydb",18701,"
[local]",59f1762a.490d,31,"DELETE",2017-10-26 13:44:10
CST,5/16770,0,ERROR,0A000,"cannot delete from foreign table
""ft_file1""",,,,,,"DELETE FROM ft_file1 ;",,,"psql"
...
```

□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□

```
CREATE FOREIGN TABLE ft_pglog (
    log_time timestamp(0) without time zone,
    user_name text,
    database_name text,
    process_id integer,
    connection_from text,
    session_id text,
    session_line_num bigint,
    command_tag text,
    session_start_time timestamp with time zone,
    virtual_transaction_id text,
    transaction_id bigint,
    error_severity text,
    sql_state_code text,
    message text,
    detail text,
    hint text,
    internal_query text,
    internal_query_pos integer,
    context text,
    query text,
    query_pos integer,
    location text,
    application_name text
) SERVER fs_file
OPTIONS ( filename
'/database/pg10/pg_root/pg_log/postgresql-2017-10-
26_000000.csv', format 'csv' );
```

[illegible]OPTIONS[illegible]csv[illegible]

```
mydb=# SELECT * FROM ft_pglog WHERE error_severity='ERROR'
AND command_tag='DELETE';
-[ RECORD 1 ]----------+------------------------------------
--------
log_time               | 2017-10-26 14:29:16
user_name              | postgres
database_name          | mydb
process_id             | 18701
```

```
connection_from            | [local]
session_id                 | 59f1762a.490d
session_line_num           | 31
command_tag                | DELETE
session_start_time         | 2017-10-27 03:44:10+08
virtual_transaction_id     | 5/16770
transaction_id             | 0
error_severity             | ERROR
sql_state_code             | 0A000
message                    | cannot delete from foreign table
"ft_file1"
detail                     |
hint                       |
internal_query             |
internal_query_pos         |
context                    |
query                      | DELETE FROM ft_file1 ;
query_pos                  |
location                   |
application_name           | psql
```

□□□□□postgresql-2017-10-26_000000.csv□□□□□□□□□□□□□□□□□□ft_pglog□□□□□□□□□□□□□□

16.6 postgres_fdw

□□□□□□file_fdw□□□□□□□□PostgreSQL□□□□□□□csv□□□□□□□□□□□□postgres_fdw□□□□□□□□□□□□PostgreSQL□□□□□□□Oracle□dblink□

16.6.1 postgres_fdw□□

□□□□□postgres_fdw□□□□□□□□□□16-2□

□16-2 postgres_fdw□□□□

角色	主机名	IP	端口	库名	用户名	版本
本地库	pghost1	192.168.28.74	1921	mydb	pguser	PostgreSQL10
远程库	pghost3	192.168.28.76	1923	des	pguser	PostgreSQL10

□□□pghost1□□mydb□□□postgres_fdw□□□□□□□pghost1□□□□mydb□□□□□□□□□□pghost3□□□des□□□□□

postgres_fdw□□□□□□□□□□□□□

1□□□postgres_fdw□□□□□

2□□□foreign server□□□□□□□□□□□□□□□□□□□□□□□□

3[illegible]

4[illegible]

[illegible]pghost1[illegible]postgres[illegible]mydb[illegible]postgres_fdw[illegible]

```
[postgres@pghost1 ~]$ psql mydb postgres
psql (10.0)
Type "help" for help.

mydb=# CREATE EXTENSION postgres_fdw;
CREATE EXTENSION
```

[illegible]postgres_fdw[illegible]pguser[illegible]postgres_fdw[illegible]pguser[illegible]postgres_fdw[illegible]

```
mydb=# GRANT USAGE ON FOREIGN DATA WRAPPER postgres_fdw TO
pguser;
GRANT
```

[illegible]pguser[illegible]mydb[illegible]

```
mydb=> \c mydb pguser
You are now connected to database "mydb" as user "pguser".

mydb=> CREATE SERVER fs_postgres_pghost3
    FOREIGN DATA WRAPPER postgres_fdw OPTIONS (host
```

```
'pghost3', port '1923', dbname 'des');
CREATE SERVER
```

[illegible]fs_postgres_pghost3[illegible]OPTIONS[illegible]PostgreSQL[illegible]

[illegible]

```
mydb=> CREATE USER MAPPING FOR pguser
    SERVER fs_postgres_pghost3 OPTIONS (user 'pguser',
password 'pguser');
CREATE USER MAPPING
```

FOR[illegible]OPTIONS[illegible]PostgreSQL[illegible]PostgreSQL[illegible]IP[illegible]PostgreSQL[illegible]

[illegible]pghost4[illegible]des[illegible]

```
[des@pghost3 ~]$ psql des pguser
des=> CREATE TABLE t_fdw1 (id int4,info text);
CREATE TABLE

des=> INSERT INTO t_fdw1 (id, info ) VALUES (1,'a'),
('2','b');
INSERT 0 2
```

□pghost1□□□□□□□□□□□□

```
mydb=> CREATE FOREIGN TABLE ft_t_fdw1 (
    id    int4,
    info  text
) SERVER fs_postgres_pghost3 OPTIONS (schema_name 'pguser',
table_name 't_fdw1');
```

OPTIONS□□□□schema_name□□□□□□□□□table_name□□□□□□□□

□□□□□□postgres_fdw□□□□□□□□□□□□ft_t_fdw1□□□□□□

```
mydb=> SELECT * FROM ft_t_fdw1 ;
ERROR:  could not connect to server "fs_postgres_pghost3"
DETAIL:  FATAL:  no pg_hba.conf entry for host
"20.26.28.74", user "pguser", database "des"
```

□□□□□pg_hba.conf□□□□□□□□□□□□□pghost3□$PGDATA/pg_hba.conf□□□□□□□□□□□□□□□□□

```
host    des     pguser  20.26.28.0/24   md5
```

□□□□pg_ctl reload□□□□□□□□

pghost1□□□□□□□□ft_t_fdw1□□□□□□□□

```
mydb=> SELECT * FROM ft_t_fdw1 ;
    id | info
-------+------
     1 | a
     2 | b
(2 rows)
```

□□□□pghost1□□□□mydb□□□□□□□□□□
pghost3□des□□□□□□

16.6.2 postgres_fdw□□□□□□□□

postgres_fdw□□□□□□□□□□□
PostgreSQL9.3□□□□□□□□□□□□□postgres_fdw
□□□□□□□□

□pghost1□□□□□□□□□□□□□□□□□□□

```
mydb=> INSERT INTO ft_t_fdw1 (id,info) VALUES (3,'c');
INSERT 0 1
mydb=> SELECT * FROM ft_t_fdw1 WHERE id=3;
   id | info
------+------
    3 | c
(1 row)
```

□pghost3□□□des□□□□□□□□□□□

```
des=> SELECT * FROM t_fdw1 WHERE ID=3;
    id | info
```

```
-------+------
     3 | c
(1 row)
```

[illegible]des[illegible]

[illegible]pghost1[illegible]

```
mydb=> UPDATE ft_t_fdw1 SET info='ccc' WHERE ID=3;
UPDATE 1
mydb=> SELECT * FROM ft_t_fdw1 WHERE id=3;
    id | info
-------+------
     3 | ccc
(1 row)
```

[illegible]pghost3[illegible]des[illegible]

```
des=> SELECT * FROM t_fdw1 WHERE ID=3;
    id | info
-------+------
     3 | ccc
(1 row
```

[illegible]

[illegible]DELETE[illegible]pghost1[illegible]DELETE[illegible]

```
mydb=> DELETE FROM ft_t_fdw1 WHERE id=3;
DELETE 1
```

□pghost3□□□des□□□□□□□□□□□

```
des=> SELECT * FROM t_fdw1 WHERE ID=3;
    id | info
-------+------
(0 rows)
```

□□□□□□□□

□□ postgres_fdw□□□PostgreSQL□□□□□□9.3□□□□

16.6.3 postgres_fdw□□□□□□□□

PostgreSQL10□□□postgres_fdw□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□10□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□postgres_fdw□□□□□□□□□□□□□□□□□10□□□□□□□□

□□□□□PostgreSQL10□□□□□□

□pghost3□□□des□□□□□□□□□□□□□□□□□□□□□

```
[des@pghost3 ~]$ psql des pguser
psql (10.0)
Type "help" for help.

des=> CREATE TABLE t_fdw2(id int4,flag int4);
CREATE TABLE

des=> INSERT INTO t_fdw2(id,flag) SELECT n,mod(n,3) FROM
generate_series(1,100000) n;
INSERT 0 100000
```

□pghost1□□□mydb□□□□□□□□□□□□

```
mydb=> CREATE FOREIGN TABLE ft_t_fdw2 (
    id    int4,
    flag  int4
) SERVER fs_postgres_pghost3 OPTIONS (schema_name 'pguser',
table_name 't_fdw2');
```

□□□□□□□□□SQL□

```
mydb=> SELECT flag,count(*) FROM ft_t_fdw2 GROUP BY flag
ORDER BY flag;
    flag | count
---------+-------
       0 | 33333
       1 | 33334
       2 | 33333
(3 rows)
```

□□□□□□□□□

```
mydb=> EXPLAIN (ANALYZE on,VERBOSE on)
    SELECT flag,count(*) FROM ft_t_fdw2 GROUP BY flag ORDER
BY flag;
QUERY PLAN
------------------------------------------------------------
--------------------
    Sort  (cost=167.52..168.02 rows=200 width=12) (actual
time=18.888..18.888 rows=3 loops=1)
        Output: flag, (count(*))
        Sort Key: ft_t_fdw2.flag
        Sort Method: quicksort  Memory: 25kB
        ->  Foreign Scan  (cost=114.62..159.88 rows=200
width=12) (actual time= 18.877..18.878 rows=3 loops=1)
                Output: flag, (count(*))
                Relations: Aggregate on (pguser.ft_t_fdw2)
                Remote SQL: SELECT flag, count(*) FROM
pguser.t_fdw2 GROUP BY flag
    Planning time: 0.109 ms
    Execution time: 19.496 ms
(10 rows)
```

□□□□□□□□□□□□□□□□□□□□□□□□□□

·Remote SQL□□□□□□□□SQL□□SQL□□□□□SQL□

·Relation Aggregate□□□□□□□□□□□□□

·Foreign Scan□“rows=3”□□Foreign Scan□□

·Sort□□□□

□□SQL□□□□□19.496ms□□□□□□□□□□□□□□□□□□□□SQL□□

```
Remote SQL: SELECT flag, count(*) FROM pguser.t_fdw2 GROUP
BY flag
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□PostgreSQL9.6□□□□□□

□□□9.6□□□□□□□□□□□□□16-3□

□16-3　postgres_fdw□□□□

角　色	主机名	IP	端　口	库　名	用户名	版　本
本地库	pghost1	192.168.28.74	1922	mydb	pguser	PostgreSQL9.6
远程库	pghost3	192.168.28.76	1924	des	pguser	PostgreSQL9.6

PostgreSQL9.6□□□□□postgres_fdw□□□□16.6.1□□□□□□□□□□□□□□□□□□□□

```
mydb=> EXPLAIN (ANALYZE on,VERBOSE on)
        SELECT flag,count(*) FROM ft_t_fdw2 GROUP BY flag
ORDER BY flag;
QUERY PLAN
------------------------------------------------------------
--------------------
```

```
Sort  (cost=324.43..324.93 rows=200 width=12) (actual
time=354.879..354.880 rows=3 loops=1)
    Output: flag, (count(*))
    Sort Key: ft_t_fdw2.flag
    Sort Method: quicksort  Memory: 25kB
    ->  HashAggregate  (cost=314.78..316.78 rows=200
width=12) (actual time= 354.863..354.863 rows=3 loops=1)
        Output: flag, count(*)
        Group Key: ft_t_fdw2.flag
        ->  Foreign Scan on pguser.ft_t_fdw2
(cost=100.00..285.53 rows=5851 width=4) (actual
time=1.014..333.186 rows=100000 loops=1)
            Output: id, flag
            Remote SQL: SELECT flag FROM pguser.t_fdw2
    Planning time: 0.152 ms
    Execution time: 355.649 ms
(12 rows)
```

□□□□□□□□□□□355□□□□□PostgreSQL10□□□□17□□□□□□□□□□□□□□□Remote SQL□□□SQL□□

```
Remote SQL: SELECT flag FROM pguser.t_fdw2
```

□□□□□□□□Foreign Scan □rows=100000□□□□□□□□□□□□□□□□□□□□pghost1□□□□pghost1□□□□□□□□□□

□PostgreSQL10□□□SQL□□□□□Remote SQL□□□SQL□□

```
Remote SQL: SELECT flag, count(*) FROM pguser.t_fdw2 GROUP
BY flag
```

[illegible]PostgreSQL10[illegible]
postgres_fdw[illegible]

16.7 Citus

Citus[illegible]B2B[illegible]Citus[illegible]PostgreSQL[illegible]pg_shard[illegible]Citus[illegible]

16.7.1 Citus[illegible]

1.[illegible]

[illegible]Citus[illegible]Citus[illegible]

2.[illegible]

Citus[illegible]Multi-tenant Application[illegible]Real-time Analytics[illegible]

3.[illegible]

[illegible]B2B[illegible]
[illegible]Citus[illegible]

4.[illegible]

[illegible]

16.7.2 Citus[illegible]

[illegible]yum[illegible]

```
# [illegible]citus[illegible]yum[illegible]
curl https://install.citusdata.com/community/rpm.sh | sudo
bash
sudo yum install -y citus72_10
```

[illegible]

```
sudo su - postgres
export PATH=$PATH:/usr/pgsql-10/bin
```

□□□□□□□□□□□□

```
mkdir -p /pgdata/citus_cluster/coordinator
/pgdata/citus_cluster/worker1 /pgdata/citus_cluster/worker2
```

□□□□□□□□□□□□□

```
/usr/pgsql-10/bin/initdb -D
/pgdata/citus_cluster/coordinator/
/usr/pgsql-10/bin/initdb -D /pgdata/citus_cluster/worker1/
/usr/pgsql-10/bin/initdb -D /pgdata/citus_cluster/worker2/
```

□□postgresql.conf□□□□□□□□

```
-- shared_preload_libraries = 'citus'
echo "shared_preload_libraries = 'citus'" >>
/pgdata/citus_cluster/coordinator/postgresql.conf
echo "shared_preload_libraries = 'citus'" >>
/pgdata/citus_cluster/worker1/postgresql.conf
echo "shared_preload_libraries = 'citus'" >>
/pgdata/citus_cluster/worker2/postgresql.conf
```

□□□□□□□□□□□□□shared preload libraries□□□□□□□□□□citus□□□□□□□□□□□□□□□□□

```
FATAL: Citus has to be loaded first
HINT: Place citus at the beginning of
```

```
shared_preload_libraries.
```

[illegible]

```
/usr/pgsql-10/bin/pg_ctl -D
/pgdata/citus_cluster/coordinator -o "-p 9700" -l
coordinator_logfile start
/usr/pgsql-10/bin/pg_ctl -D /pgdata/citus_cluster/worker1 -o
"-p 9701" -l worker1_logfile start
/usr/pgsql-10/bin/pg_ctl -D /pgdata/citus_cluster/worker2 -o
"-p 9702" -l worker2_logfile start
```

[illegible]Coordinator[illegible]

```
/usr/pgsql-10/bin/psql -p 9700 postgres -c "CREATE DATABASE
mydb;"
NOTICE:  Citus partially supports CREATE DATABASE for
distributed databases
DETAIL:  Citus does not propagate CREATE DATABASE command to
workers
HINT:  You can manually create a database and its extensions
on workers.
CREATE DATABASE
```

[illegible]Coordinator[illegible]Coordinator[illegible]Worker[illegible]Worker[illegible]Citus Extension[illegible]

[illegible]Worker[illegible]

```
/usr/pgsql-10/bin/psql -p 9701 postgres -c "CREATE DATABASE
mydb;"
/usr/pgsql-10/bin/psql -p 9702 postgres -c "CREATE DATABASE
mydb;"
```

□□□□□□□□Citus Extension□□□□□□

```
/usr/pgsql-10/bin/psql -p 9700 mydb -c "CREATE EXTENSION
citus;"
/usr/pgsql-10/bin/psql -p 9701 mydb -c "CREATE EXTENSION
citus;"
/usr/pgsql-10/bin/psql -p 9702 mydb -c "CREATE EXTENSION
citus;"
```

16.7.3 Citus□□

□□□□□Citus□□□□□Work□□□□□Work□□□□□Work□□□□□Work□□□

□Citus□□□□□Worker□□□□□□□□

```
/usr/pgsql-10/bin/psql -p 9700 mydb -c "SELECT * FROM
master_add_node('127. 0.0.1', 9701);"
   nodeid | groupid | nodename  | nodeport | noderack |
hasmetadata | isactive
----------+---------+-----------+----------+----------+-----
--------+----------
        1 |       1 | 127.0.0.1 |     9701 | default  | f
| t
(1 row)
/usr/pgsql-10/bin/psql -p 9700 mydb -c "SELECT * FROM
master_add_node('127. 0.0.1', 9702);"
   nodeid | groupid | nodename  | nodeport | noderack |
```

```
hasmetadata | isactive
----------+---------+----------+---------+---------+-----
--------+----------
        2 |       2 | 127.0.0.1 |     9702 | default  | f
| t
(1 row)
```

□Citus□□□□□Worker□□□□□□□□

```
UPDATE pg_dist_shard_placement set shardstate = 3 where
nodename = '127.0.0.1' and nodeport = 9702;
SELECT master_remove_node('127.0.0.1', 9702);
    master_remove_node
--------------------
(1 row)
SELECT * FROM master_get_active_worker_nodes();
    node_name | node_port u
-------------+----------
    127.0.0.1 |      9701
(1 row)
```

□□Worker□□□□□□□□

```
SELECT master_disable_node('127.0.0.1','9702');
NOTICE:  Node 127.0.0.1:9702 has active shard placements.
Some queries may fail after this operation. Use SELECT
master_activate_node('127.0.0.1', 9702) to activate this
node back.
    master_disable_node
---------------------
(1 row)
```

□□Worker□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□

```
ERROR: failed to assign 2 task(s) to worker nodes
```

□□□□□□□Worker□□□□□□□□Worker□□□□□Worker□□□□□□□□

```
SELECT master_activate_node('127.0.0.1', 9702);
```

□Coordinator□□□□Worker□□□□□□□□

```
/usr/pgsql-10/bin/psql -p 9700 mydb -c " SELECT * FROM
master_get_active_worker_nodes();"
    node_name | node_port
--------------+-----------
    127.0.0.1 |      9701
    127.0.0.1 |      9702
(2 rows)
```

16.7.4 □□□□□

□Coordinator□□□□□□□□□□

```
CREATE TABLE table_name (
    user_id integer NOT NULL,
    name character varying(32)
);
```

□Coordinator□□□□□□□□□□□

```
mydb=# SELECT create_distributed_table('table_name',
'user_id');
    create_distributed_table
--------------------------
(1 row)
```

[illegible]Citus[illegible]32[illegible]Citus[illegible]CPU[illegible]×[illegible]×[illegible]

```
set citus.shard_count = 64;
```

[illegible]Worker[illegible]join[illegible]Coordinator[illegible]create_reference_table[illegible]Worker[illegible]

```
CREATE TABLE rt(id serial primary key,ival int);
SELECT create_reference_table('rt');
```

[illegible]1[illegible]upgrade_to_reference_table[illegible]API[illegible]

16.7.5 Citus□□□□

·citus.shard_replication_factor □integer□

□□□□□□□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□2□□□□□□□□□□□□□□□□□□□□

·citus.shard_count □integer□

□□□□□□□□□□□□32□□□Citus□□□□□Worker□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□3□□□□Worker1□□□2□□□□□Worker2□□□□□1□□□□

·citus.task_executor_type□enum□

□□□□□□□□□□□□□□□real-time□task-tracker□□□□real-time□real-time□□□task-tracker□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□worker□□□□□□□□□□□□□□□□□□□

□Citus□□□□□□□□□□□postgresql.conf□□idle_in_transaction_session_timeout□□□

[illegible]Citus[illegible]Worker[illegible]

16.7.6 Citus[illegible]

[illegible]

```
SELECT
pg_size_pretty(citus_total_relation_size('table_name'::regcl
ass));
```

[illegible]

```
SELECT
pg_size_pretty(citus_relation_size('table_name'::regclass));
-- [illegible]
SELECT
pg_size_pretty(citus_table_size('table_name'::regclass));
```

[illegible]

```
SELECT column_to_column_name(logicalrelid, partkey) AS dk
FROM pg_dist_partition
WHERE logicalrelid='table_name'::regclass;
```

[illegible]

```
SELECT get_shard_id_for_distribution_column('table_name',
DK_value);
```

□□□□□□□□□□create_distributed_table□□□□□□

```
mydb=# \df+ create_distributed_table
List of functions
-[ RECORD 1 ]-------+-------------------------------------
Schema              | pg_catalog
Name                | create_distributed_table
Result data type    | void
Argument data types | table_name regclass,
distribution_column text, distribution_type
citus.distribution_type DEFAULT
'hash'::citus.distribution_type, colocate_with text DEFAULT
'default'::text
Type                | normal
Volatility          | volatile
Parallel            | unsafe
Owner               | postgres
Security            | invoker
Access privileges   |
Language            | c
Source code         | create_distributed_table
Description         | creates a distributed table
```

□□□□□□□□□□□□

```
mydb=# \dT+ citus.distribution_type
List of data types
-[ RECORD 1 ]-----+-----------------------
Schema            | citus
Name              | citus.distribution_type
Internal name     | distribution_type
Size              | 4
```

```
Elements          | hash                    +
                  | range                   +
                  | append
Owner             | postgres
Access privileges |
Description       |
```

□□□□□□Citus□□□

```
SET citus.explain_all_tasks = 'TRUE';
```

DDL□□□□□□Citus□□□ALTER TABLE ADD COLUMN□ALTER COLUMN□DROP COLUMN□ALTER COLUMN TYPE□RENAME COLUMN□□□□□□□□PostgreSQL□□□□□□□□□□□

```
ALTER TABLE tbl ADD COLUMN col INT;
NOTICE: using one-phase commit for distributed DDL commands
HINT: You can enable two-phase commit for extra safety with:
SET citus.multi_shard_commit_protocol TO '2pc'
```

□□□□□□Citus□□□□□□□□□□□DDL□□□□□□□□□□□□□□□□□□DDL□□□□□

```
SET citus.multi_shard_commit_protocol TO '2pc';
ALTER TABLE tbl ADD COLUMN col INT;
ALTER TABLE test ALTER COLUMN ival3 SET DEFAULT 1;
```

[illegible]PostgreSQL[illegible]Citus[illegible]

```
ERROR: creating index without a name on a distributed table
is currently unsupported
```

[illegible]PostgreSQL[illegible]

```
CREATE INDEX CONCURRENTLY idx_name ON table_name
(col1...col_n);
```

[illegible]citus[illegible]

[illegible]

```
DROP INDEX umch_user_id_name;
NOTICE: using one-phase commit for distributed DDL commands
HINT: You can enable two-phase commit for extra safety with:
SET citus.multi_shard_commit_protocol TO '2pc'
```

[illegible]VACUUM[illegible]ANALYZE[illegible]

```
VACUUM table_name;
ANALYZE table_name;
```

[illegible]Citus[illegible]VERBOSE[illegible]VACUUM[illegible]

```
ERROR: the VERBOSE option is currently unsupported in
distributed VACUUM commands
```

[illegible]Citus[illegible]PostgreSQL Extensions[illegible]PostGIS[illegible]hll[illegible]jsonb[illegible]

·[illegible]shared_preload_libraries[illegible]citus[illegible]

·[illegible]Coordinator[illegible]Worker[illegible]Citus[illegible]

[illegible]

```
yum --showduplicates list citus72_10
Available Packages
citus72_10.x86_64   7.2.0.citus-1.el6   citusdata_community
citus72_10.x86_64   7.2.1.citus-1.el6   citusdata_community
```

[illegible]

```
yum install -y 7.2.1.citus-1.el6
```

[illegible]PostgreSQL[illegible]

[illegible]yum[illegible]

16.8 [illegible]

PostgreSQL[illegible] PostgreSQL[illegible]

[illegible]PostgreSQL[illegible]GitHub[illegible]

[illegible]17[illegible] Oracle[illegible]PostgreSQL[illegible]

[illegible]SQL/MED[illegible]PostgreSQL[illegible]PostgreSQL[illegible]PostgreSQL[illegible]Oracle[illegible]PostgreSQL[illegible]

[illegible]Oracle[illegible]PostgreSQL[illegible]PostgreSQL[illegible]oracle_fdw[illegible]Oracle[illegible]30[illegible]4[illegible]30G[illegible]100GB[illegible]PostgreSQL9.2[illegible]9.3[illegible]

[illegible]Oracle[illegible]PostgreSQL[illegible]PostgreSQL[illegible]

17.1 [illegible]

[illegible]

·[illegible]

·[illegible]SQL[illegible]PostgreSQL[illegible]Oracle[illegible]SQL[illegible]SQL[illegible]

·[illegible]Oracle[illegible]PostgreSQL[illegible]Oracle[illegible]

·[illegible]DBA[illegible]

·[illegible]DBA[illegible]

·[illegible]

[illegible]DBA[illegible]

17.2 [illegible]

Oracle[illegible]PostgreSQL[illegible]Oracle[illegible]PostgreSQL[illegible]

1.[illegible]

[illegible]Oracle[illegible]PostgreSQL[illegible]17-1[illegible]

[illegible]17-1 Oracle[illegible]PostgreSQL[illegible]

数据类型	Oracle 11g	PostgreSQL 10
字符类型	VARCHAR、VARCHAR2、N VARCHAR、N VARCHAR2	character varying 或 text
	CHAR、NCHAR	character
	CLOB、NCLOB、LONG	text
数字类型	NUMBER	numeric
	FLOAT	real、double precision
时间日期类型	DATE	date 或 timestamp
	TIMESTAMP	timestamp
二进制类型	BLOB、RAW	bytea

[illegible]Oracle[illegible]Oracle[illegible]PostgreSQL[illegible]2[illegible]

[illegible]17-1[illegible]PostgreSQL[illegible]DBA[illegible]DBA[illegible]

[illegible]Oracle[illegible]PostgreSQL[illegible]PostgreSQL[illegible]

2.[illegible]

[illegible]PostgreSQL[illegible]PostgreSQL[illegible]Oracle[illegible]Oracle[illegible]PostgreSQL[illegible]

[illegible]PostgreSQL[illegible]https://www.postgresql.org/docs/10/static/plpgsql.html [illegible]

17.3 [illegible]

[illegible]SQL[illegible]SQL[illegible]SQL[illegible]SQL[illegible]

[illegible]SQL[illegible]SELECT[illegible]UPDATE[illegible]INSERT[illegible]DELETE[illegible]SQL[illegible]SQL[illegible]Oracle[illegible]SQL[illegible]

1.SQL[illegible]

[illegible]SQL[illegible]PostgreSQL[illegible]Oracle[illegible]SQL[illegible]

Oracle[illegible]ROWNUM[illegible]

```
SQL> SELECT OBJECT_NAME FROM dba_objects WHERE ROWNUM < 2;
OBJECT_NAME
------------------------------------------------------------
--------------------
ICOL$
```

PostgreSQL□□□□LIMIT□□□□□□□□□□□□□□
□□□

```
postgres=# SELECT relname FROM pg_class LIMIT 1;
 relname
---------
 test_sr
(1 row)
```

Oracle□□ROWNUM□PostgreSQL□LIMIT□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□ROWNUM
□□□□□□□□LIMIT□□□□□□

Oracle□□ROWNUM□PostgreSQL□LIMIT□
□□□□□□□□□□

SQL□□□□□□□□□□□□□□□□□□□□Oracle□
PostgreSQL□□□□□□□□□□□□□□□□□□Oracle□□
□□SQL□□□□□□□□□□

```
SQL> SELECT SEQ_1.CURRVAL FROM DUAL;
 CURRVAL
----------
 2
```

□PostgreSQL□□□□□□□□□□□□□□□□□

```
postgres=# SELECT currval('seq_1');
 currval
---------
 2
(1 row
```

currval[illegible]nextval[illegible]PostgreSQL[illegible]Oracle[illegible]

[illegible]Oracle[illegible]PostgreSQL[illegible]Oracle[illegible]

```
SQL> SELECT * FROM (SELECT * FROM table_1);
```

[illegible]PostgreSQL[illegible]

```
mydb=> SELECT * FROM (SELECT * FROM table_1) as b;
```

SQL[illegible]Oracle[illegible]START WITH...CONNECT BY[illegible]PostgreSQL[illegible]

[illegible]Oracle[illegible]t_area[illegible]

```
CREATE TABLE t_area(id numeric,name varchar(32),fatherid
numeric);

INSERT INTO t_area VALUES (1, '□□'   ,0);
INSERT INTO t_area VALUES (2, '□□'   ,1);
INSERT INTO t_area VALUES (3, '□□'   ,1);
INSERT INTO t_area VALUES (4, '□□'   ,2);
INSERT INTO t_area VALUES (5, '□□'   ,2);
INSERT INTO t_area VALUES (6, '□□'   ,3);
INSERT INTO t_area VALUES (7, '□□□' ,4);
INSERT INTO t_area VALUES (8, '□□□' ,4);
```

□□ID□□4□□□□□□□□□□□□□□□

```
SQL> SELECT id, name, fatherid
    2  FROM t_area
    3  START WITH id = 4
    4  CONNECT BY PRIOR id = fatherid;

        id name                             fatherid
---------- -------------------------------- ----------
         4 □□                                2
         7 □□□                              4
         8 □□□                              4
```

PostgreSQL□□□□□□□□□□WITH RECURSIVE□□□□PostgreSQL□□□t_area□□□□□□□□□□□□□□□□

```
CREATE TABLE t_area(id int4,name varchar(32),fatherid int4);
```

□□id□fatherid□□□□□Oracle□□numeric□□□□PostgreSQL□□□□int4□□□□□□□□□□□□□□

PostgreSQL□mydb□□□□□□□□□□□

```
mydb=> WITH RECURSIVE r AS (
            SELECT * FROM t_area WHERE id = 4
        UNION  ALL
            SELECT t_area.* FROM t_area, r WHERE
t_area.fatherid = r.id
        )
    SELECT * FROM r ORDER BY id;
    id |  name | fatherid
-------+-------+----------
     4 | □□   |        2
     7 | □□□ |       4
     8 | □□□ |       4
(3 rows)
```

□□PostgreSQL□□□□□□4.1□□□□□□□

□□□□□□□□□Oracle□PostgreSQL□SQL□□□□□□□□□□□□□□□□□□Oracle□□PostgreSQL□□□□□□□□□□□□□□□□□

2.□□□□□

□Oracle□SQL□□□□□□□Oracle□□□□□□□□□□□SQL□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□Oracle□□SYSDATE□CURRENT_DATE□□□□□□□□□□□□□□

```
SQL> SELECT SYSDATE,CURRENT_DATE FROM DUAL;
SYSDATE      CURRENT_DATE
------------ ------------
13-11□-17   13-11□-17
```

□□□□□□□□□□□□□□□CURRENT_TIMESTAMP□□□□□

```
SQL> SELECT CURRENT_TIMESTAMP FROM DUAL;
CURRENT_TIMESTAMP
---------------------------------------------------------------
---------------
13-11□-17 02.45.42.330637 □□ +08:00
```

PostgreSQL□□□CURRENT_DATE□CURRENT_TIMESTAMP□□□CURRENT_DATE□□□□□CURRENT_TIMESTAMP□now□□□□□□□□□□□□□□

```
mydb=> SELECT CURRENT_DATE;
 current_date
--------------
 2017-11-12
(1 row)

mydb=> SELECT CURRENT_TIMESTAMP,now();
        current_timestamp         |              now
----------------------------------+-------------------------
------
    2017-11-12 11:13:37.430499+08 | 2017-11-12
11:13:37.430499+08
(1 row)
```

[illegible]Oracle[illegible]INSTR[illegible]0[illegible]INSTR[illegible]

```
{INSTR} (string , substring [, position [, occurrence]])
```

·string[illegible]

·substring[illegible]

·position[illegible]1[illegible]

·occurrence[illegible]1[illegible]

[illegible]

```
SQL> SELECT INSTR('Hello PostgreSQL!','o') FROM DUAL;
INSTR('HELLOPOSTGRESQL!','O')
-----------------------------
                            5
```

PostgreSQL[illegible]

```
position(substring in string)
```

PostgreSQL[illegible]position[illegible]Oracle[illegible]INSTR[illegible]

```
mydb=> SELECT position('o' in 'Hello PostgreSQL!');
 position
----------
 5
(1 row)
```

[illegible]PostgreSQL[illegible]Oracle[illegible]PostgreSQL[illegible]

[illegible]

17.4 □□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□Oracle□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□

·□□□□□□Oracle□□□□□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□□□□□text□□□□csv□□□□□□□□□□□□Oracle□□□□□□□□

·□□□□□□PostgreSQL□□□□□oracle_fdw□□□□□□□□□□□□□□PostgreSQL□□□□□□□□□Oracle□□□□□□□□□SQL□□□□Oracle□□□□□□□□□PostgreSQL□□□□□□□□□□□□Oracle□□□□□□□□□

·□□□□□□□□□□ETL□□□□□□□□□

□□□□□□□□□□□Oracle□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□oracle_fdw□□□□□Oracle□□□□□□PostgreSQL□□□□□□□□□17-2□□□□

□17-2 oracle_fdw□□□□

角色	主机名	IP	端口	库名	用户名	版本	字符集
本地库 PostgreSQL	pghost1	192.168.28.74	1921	mydb	pguser	PostgreSQL10	UTF8
远程库 Oracle	pghost3	192.168.28.76	1521	oradb	community	Oracle 11g	UTF8

1.pghost1□□□□□□Oracle 11g□□□□

□pghost1□□□□□oracle_fdw□□□□□Oracle□□□□□□□□RPM□□□□□□□□basic□devel□sqlplus□□□□□□□RPM□□□□□□□http://www.oracle.com/technetwork/topics/linuxx86-64soft-092277.html

□□RPM□□□□□□□

```
[root@pghost1 soft_bak]# rpm -ivh oracle-instantclient11.2-
basic-11.2.0.1.0-1.x86_64.rpm
[root@pghost1 soft_bak]# rpm -ivh oracle-instantclient11.2-
sqlplus-11.2.0.1.0-1.x86_64.rpm
[root@pghost1 soft_bak]# rpm -ivh oracle-instantclient11.2-
devel-11.2.0.1.0-1.x86_64.rpm
```

□□□□postgres□□□□□□□□□□□□Oracle□□□□□□□□□□□□□□

```
...PostgreSQL□□□□□□□□
export ORACLE_BASE=/usr/lib/oracle
export ORACLE_HOME=/usr/lib/oracle/11.2/client64
export
LD_LIBRARY_PATH=$ORACLE_HOME/lib:$PGHOME/lib:/lib64:/usr/lib
64:/usr/local/lib64:/lib:/usr/lib:/usr/local/lib
export PATH=$ORACLE_HOME/bin:$PATH:$PGHOME/bin:.
```

□□□□□□□□Oracle□□□□□□□□□□□□□□□□□□

□□□pghost1□□□□□□□□□□□□□□□pghost3□□oracle□□□□□□□□□

```
[postgres@pghost1 ~]$ sqlplus
community/community@//192.168.28.76/oradb
SQL*Plus: Release 11.2.0.1.0 Production on Tue Nov 14
20:55:56 2017
Copyright (c) 1982, 2009, Oracle.  All rights reserved.
Connected to:
Oracle Database 11g Enterprise Edition Release 11.2.0.1.0 -
64bit Production
With the Partitioning, OLAP, Data Mining and Real
Application Testing options

SQL>
```

□□□□Oracle□□□□□□□□

2.pghost1□□□□□oracle_fdw

□https://api.pgxn.org/dist/oracle_fdw □□□oracle_fdw□□□□□□□□□□□oracle_fdw 2.0.0□

□□oracle_fdw□□□□□□□

```
# unzip oracle_fdw-2.0.0.zip
```

□□□□□□□□□□□PostgreSQL□pg_config□□□□□□root□□□□□□□□□□□postgres□□□□□□□□□□□□□□□□□

```
[root@pghost1 oracle_fdw-2.0.0]# source
/home/postgres/.bash_profile
[root@pghost1 oracle_fdw-2.0.0]# which pg_config
/opt/pgsql/bin/pg_config
```

□□□□□oracle_fdw□□□□□□

```
[root@pghost1 oracle_fdw-2.0.0]# make
...
[root@pghost1 oracle_fdw-2.0.0]# make install
/bin/mkdir -p '/opt/pgsql_10.0/lib'
/bin/mkdir -p '/opt/pgsql_10.0/share/extension'
/bin/mkdir -p '/opt/pgsql_10.0/share/extension'
/bin/mkdir -p '/opt/pgsql_10.0/share/doc/extension'
/usr/bin/install -c -m 755  oracle_fdw.so
'/opt/pgsql_10.0/lib/oracle_fdw.so'
/usr/bin/install -c -m 644 .//oracle_fdw.control
'/opt/pgsql_10.0/share/extension/'
/usr/bin/install -c -m 644 .//oracle_fdw--1.1.sql
.//oracle_fdw--1.0--1.1.sql
'/opt/pgsql_10.0/share/extension/'
/usr/bin/install -c -m 644 .//README.oracle_fdw
'/opt/pgsql_10.0/share/doc/extension/'
```

□□oracle_fdw□□□□□□□□□
$PGHOME/share/extension□□□□□□□□□□□□

```
[root@pghost1 oracle_fdw-2.0.0]# ll
/opt/pgsql/share/extension/oracle_fdw*
-rw-r--r-- 1 root root  231 Nov 12 14:45
/opt/pgsql/share/extension/oracle_fdw--1.0--1.1.sql
-rw-r--r-- 1 root root 1003 Nov 12 14:45
/opt/pgsql/share/extension/oracle_fdw--1.1.sql
-rw-r--r-- 1 root root  133 Nov 12 14:45
/opt/pgsql/share/extension/oracle_fdw.control
```

□□□$PGHOME/share/extension□□□□□oracle_fdw□□□□□

3.pghost1□mydb□□□□oracle_fdw

□pghost1□□mydb□□□oracle_fdw□□□□□□□□□□

```
[postgres@pghost1 ~]$ psql mydb postgres
psql (10.0)
Type "help" for help.

mydb=# CREATE EXTENSION oracle_fdw;
CREATE EXTENSION
```

□□□□□□□□□□□□□□□□□□□□□□□oracle_fdw□□□□□□□□□□□□□pguser□□oracle_fdw□□□□□□pguser□□□oracle_fdw□□□□□□□□□□

```
mydb=# GRANT USAGE ON FOREIGN DATA WRAPPER oracle_fdw TO
pguser;
GRANT
```

4.pghost3□Oracle□□□□□□□□□□□

□□community□oracle□□□□□□□□□□□community□□□□□□□□□□□□□□□

```
CREATE TABLE T_ORA1(id numeric,info varchar2(32),create_time timestamp);

INSERT INTO T_ORA1 VALUES (1,'a',CURRENT_TIMESTAMP);
INSERT INTO T_ORA1 VALUES (2,'b',CURRENT_TIMESTAMP);
INSERT INTO T_ORA1 VALUES (3,'□□□□□',CURRENT_TIMESTAMP);
```

□□□□□□□□□□□□□□□□□□□□Oracle□□□PostgreSQL□□□READONLY□□□□□□□□□□Oracle□□□□□□□□□□

```
CREATE USER READONLY IDENTIFIED BY "readonly"
    DEFAULT TABLESPACE TS_COMMUNITY
    TEMPORARY TABLESPACE TEMP
    PROFILE DEFAULT
    ACCOUNT UNLOCK;

GRANT CONNECT TO READONLY;
GRANT SELECT ON COMMUNITY.T_ORA1 TO READONLY;
```

□□□□T_ORA1□□□□□□□READONLY□□□□□□□□□□□□□□□□□□□□□□□□□□□□

5.pghost1□mydb□□□□□□□□□□□□

□pguser□□□□mydb□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ psql mydb pguser
psql (10.0)
Type "help" for help.

mydb=> CREATE SERVER fs_oracle_pghost3 FOREIGN DATA WRAPPER
```

```
oracle_fdw OPTIONS (dbserver '//192.168.28.76/oradb');
CREATE SERVER
```

□OPTIONS□□□□dbserver□□□□Oracle□□□□□□□□□□□CREATE SERVER□□□□□□□□□Oracle□□□□□□□□□□□□□□□Oracle□□□□□□□□□□□Oracle□□□□tnsnames.ora□□□□□OPTIONS□□dbserver□□□□□□□□□□□□

□□□Oracle□□□tnsnames.ora□□□□□□Oracle□□□□□□

```
oradb=
    (DESCRIPTION=
        (ADDRESS=
            (PROTOCOL=TCP)
            (HOST=192.168.28.76)
            (PORT=1521)
        )
        (CONNECT_DATA=
            (SERVICE_NAME=oradb)
        )
    )
```

□□□□□□□□□□□□□□□□□□

```
mydb=> CREATE SERVER fs_oracle_pghost3 FOREIGN DATA WRAPPER
oracle_fdw OPTIONS (dbserver 'oradb');
```

□□□□□□dbserver□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□

```
mydb=> CREATE USER MAPPING FOR pguser SERVER
fs_oracle_pghost3 OPTIONS (user 'readonly', password
'readonly');
CREATE USER MAPPING
```

□□□□□□□□□□pguser□□□PostgreSQL□□□□□OPTIONS□□readonly□□□Oracle□□□□□□

6.pghost1□mydb□□□□□□□

mydb□□□□□□□□□□□□□Oracle□□□T_ORA1□□□□□□□□

```
mydb=> CREATE FOREIGN TABLE ft_t_ora1 (
    id    int4,
    info  text,
    create_time timestamp with time zone
) SERVER fs_oracle_pghost3 OPTIONS (schema 'COMMUNITY',
table 'T_ORA1');
```

SERVER□□□□□□□□□□□□□□□□□□□□□□fs_oracle_pghost3□□□□

OPTIONS□□□□□□□□□□□□□□□□

·Table□Oracle□□□□□□□□□□Oracle□□□□□□□□□□□□□□Oracle□□□

·Schema□□□Oracle□□□□□□□□□□□□□

·Readonly□□□□□□□□□□□□□□□□□□□□□□yes□□INSERT□UPDATE□DELETE□□□□□□□□□□□□□false□

□□□□□schema□table□□□□□□□□□□□□□https://pgxn.org/dist/oracle_fdw/ □□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT * FROM ft_t_ora1;
    id |    info    |          create_time
-------+------------+-------------------------------
     1 | a          | 2017-11-12 15:02:54.465304+08
     2 | b          | 2017-11-12 15:02:54.477453+08
     3 | □□□□□ | 2017-11-12 15:02:54.493603+08
(3 rows)
```

□□□□PostgreSQL□□□□□□□□Oracle□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

7.□Oracle□□□□□□□□□PostgreSQL□□

□□□PostgreSQL□□□□□□□□□□Oracle□□□□□□□□□□□□□□□□□

```
mydb=> CREATE TABLE t_ora1(id int4,info text,create_time
timestamp with time zone);
CREATE TABLE
```

[illegible]INSERT[illegible]

```
mydb=> INSERT INTO t_ora1 SELECT * FROM ft_t_ora1;
INSERT 0 3
```

[illegible]INSERT[illegible]

```
mydb=> SELECT * FROM t_ora1;
    id |    info     |          create_time
-------+-------------+------------------------------
     1 | a           | 2017-11-12 15:02:54.465304+08
     2 | b           | 2017-11-12 15:02:54.477453+08
     3 | [illegible] | 2017-11-12 15:02:54.493603+08
(3 rows)
```

[illegible]Oracle[illegible]T_OA1[illegible]PostgreSQL[illegible]

17.5 [illegible]

[illegible]DBA[illegible]

[illegible]SQL[illegible]DBA[illegible]DBA[illegible]SQL[illegible]SQL[illegible]

[illegible]SQL[illegible]DBA[illegible]

17.6 [illegible]

[illegible]

[illegible]DBA[illegible]

[illegible]Oracle[illegible]PostgreSQL[illegible]100GB[illegible]

17.7 oracle_fdw[illegible]

[illegible]oracle_fdw[illegible]

[illegible]oracle_fdw[illegible].so[illegible]

[illegible]oracle_fdw[illegible]

```
mydb=# CREATE EXTENSION oracle_fdw;
ERROR:  could not load library
"/opt/pgsql_10.0/lib/oracle_fdw.so": libclntsh.so.11.1:
cannot open shared object file: No such file or directory
```

[illegible]lib[illegible]lib[illegible]Oracle[illegible]Oracle[illegible]$LD_LIBRARY_PATH[illegible]$ORACLE_HOME[illegible]lib[illegible]$PGHOME/lib[illegible]

```
[root@pghost1 ~]# cp
/usr/lib/oracle/11.2/client64/lib/libclntsh.so.11.1
/opt/pgsql/lib/
[root@pghost1 ~]# cp
/usr/lib/oracle/11.2/client64/lib/libnnz11.so
/opt/pgsql/lib/
```

[illegible]oracle_fdw[illegible]

□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□

```
mydb=> SELECT * FROM ft_t_ora1;
ERROR:  error connecting to Oracle: OCIEnvCreate failed to
create environment handle
DETAIL:
```

□□□□□□□□□□□□□□□□Oracle□□□□□□□□□$ORACLE_HOME□$LD_LIBRARY_PATH□□□□□□□Oracle□□□□□□□□□□□□

```
...□□PostgreSQL□□□□□□
export ORACLE_BASE=/usr/lib/oracle
export ORACLE_HOME=/usr/lib/oracle/11.2/client64
export
LD_LIBRARY_PATH=$ORACLE_HOME/lib:$PGHOME/lib:/lib64:/usr/lib
64:/usr/local/lib64:/lib:/usr/lib:/usr/local/lib
export PATH=$ORACLE_HOME/bin:$PATH:$PGHOME/bin:.
```

□□□□□□□□□□sqlplus□□□□□□□□Oracle□□□□□□□

```
sqlplus community/community@//192.168.28.76/oradb
```

□□sqlplus□□□□□□□Oracle□□□□□□Oracle□□□□□□□□□□□□PostgreSQL□□□□□□□□□□□□

□□□□□

□□□□□□□oracle_fdw□□□□□□□□□□□□□□□□□□□□□□□□□□□Oracle□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□□□□

□□□□□□□□□□□□□□□Oracle□□□□

□□oracle_fdw□□□□OPTIONS□schema□table□□□□□□Oracle□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□ft_t_ora1□□□□□□□□□□□□□□□□□□□

```
CREATE FOREIGN TABLE ft_t_ora2 (
    id    int4,
    info  text,
    create_time timestamp with time zone
) SERVER fs_oracle_pghost3 OPTIONS (schema 'COMMUNITY',
table 't_ora1');
```

□□□□□ft_t_ora2□□□□□□

```
mydb=> SElECT * FROM ft_t_ora2;
ERROR:  Oracle table "COMMUNITY"."t_ora1" for foreign table
"ft_t_ora2" does not exist or does not allow read access
DETAIL:  ORA-00942: table or view does not exist
HINT:  Oracle table names are case sensitive (normally all
uppercase)
```

[illegible]ft_t_ora2[illegible]Oracle[illegible]

[illegible]schema[illegible]ft_t_ora3[illegible]ft_t_ora1[illegible]schema[illegible]

```
CREATE FOREIGN TABLE ft_t_ora3 (
    id    int4,
    info  text,
    create_time timestamp with time zone
) SERVER fs_oracle_pghost3 OPTIONS (schema 'community',
table 'T_ORA1');
```

[illegible]ft_t_ora3[illegible]

```
mydb=> SElECT * FROM ft_t_ora3;
ERROR:  Oracle table "community"."T_ORA1" for foreign table
"ft_t_ora3" does not exist or does not allow read access
DETAIL:  ORA-00942: table or view does not exist
HINT:  Oracle table names are case sensitive (normally all
uppercase).
```

[illegible]ft_t_ora3[illegible]

17.8 [illegible]

[illegible]Oracle[illegible]PostgreSQL[illegible]Oracle[illegible]PostgreSQL[illegible]

[illegible]oracle_fdw[illegible]Oracle[illegible]PostgreSQL[illegible]Ora2Pg[illegible]Oracle[illegible]PostgreSQL[illegible]Oracle[illegible]Oracle[illegible]PostgreSQL[illegible]SQL[illegible]PostgreSQL[illegible]Oracle[illegible]SQL[illegible]http://ora2pg.darold.net/ [illegible]

[illegible]18[illegible] PostGIS

[illegible]DBMS[illegible]

PostgreSQL[illegible]PostgreSQL[illegible]PostGIS[illegible]GIS[illegible]

[illegible]PostGIS[illegible]

18.1 □□□□□

□□□□□□□□PostGIS□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□yum□apt-get□brew□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□PostgreSQL□PostGIS□□□□□□□□□□□□□□□yum□□□□□□□□□□□□□□□□□□□□□□□□

```
[postgres@pghost1 ~]$ yum search postgis
Loaded plugins: fastestmirror, security
...
...
...
postgis24_10.x86_64 : Geographic Information Systems
Extensions to PostgreSQL
    Name and summary matches only, use "search all" for
everything.
```

□□□□PostgreSQL 10□□□□□PostGIS□□□□□□

```
[postgres@pghost1 ~]$ yum install -y postgis24_10
```

□□□□□□□□□□□□□□□

18.2 □□GIS□□□

PostGIS□PostgreSQL□□□□□□□□□□□□□□□□□□geo□□□□□□PostGIS□□□

```
postgres=# CREATE DATABASE geo;
CREATE DATABASE
postgres=# \c geo
geo=# CREATE EXTENSION postgis;
CREATE EXTENSION
```

□□PostgreSQL□□□□□□□□□□PostGIS□□□□□□□□□□□□□□□□□□□

```
geo=# SELECT name,default_version FROM
pg_available_extensions WHERE name ~ 'gis';
          name            | default_version
--------------------------+-----------------
    postgis               | 2.4.3
    postgis_tiger_geocoder | 2.4.3
    postgis_topology      | 2.4.3
    postgis_sfcgal        | 2.4.3
```

□□□□□□□□postgis_full_version□□□□PostGIS□□□□□□□□

```
gis=# SELECT postgis_full_version();
POSTGIS="2.4.3 r16312" PGSQL="100" GEOS="3.6.2-CAPI-1.10.2
4d2925d6" PROJ="Rel. 4.9.3, 15 August 2016" GDAL="GDAL
```

```
1.11.5, released 2016/07/01" LIBXML="2.9.7" LIBJSON="0.12.1"
RASTER
```

□□GIS□□□□□□□□□□□□□□□□□□

18.3 [illegible]

PostGIS[illegible]
[illegible]

[illegible] [illegible]PostGIS[illegible]PostgreSQL[illegible]PostGIS[illegible]ST[illegible]Spatial Type[illegible]

[illegible]PostGIS[illegible]Geometry[illegible]PostGIS[illegible]geos[illegible]

18.3.1 [illegible]

PostGIS[illegible]

·[illegible]WKT[illegible]Well-Known-Text[illegible]

·[illegible]WKB[illegible]Well-Known-Binary[illegible]

·GeoJSON[illegible]

·□□□□□□□□□

□□□□□□□□□1□2□□□□□□□□□□□□□□□□□□□□□□□□

```
gis=# SELECT
    'Point(1 2)'::GEOMETRY
AS wkt,
    '0101000000000000000000F03F0000000000000040'::GEOMETRY
AS wkb,
    ST_GeomFromGeoJSON('{"type":"Point","coordinates":
[1,2]}') AS geo_json,
    ST_Point(1, 2)
AS func;
```

18.3.2 □□□□□□□

PostGIS□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□

□□□□□□□□□□Point□ST_Point□□□□□□□□□□□□□□□□

```
geo=# SELECT Point(1,2), ST_Point(1,2);
    point |                  st_point
---------+--------------------------------------------
    (1,2) | 0101000000000000000000F03F0000000000000040
```

PostgreSQL□□Point□□□□□□□□Double□□□□□16□□□□□PostGIS□□□□ST_Point□□□□□□□□□□□□21□□ID□□

□□□PostGIS□□□□□□□□PostgreSQL□□□□□□□□□□□□□□□□□□□□□□□□□□ETL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□pg_dump□□□□□□□□□□□□

□□□□□□□□□□□□□□□□ST_AsText□□WKT□□□WKB□□□□□□

```
geo=# SELECT ST_AsText(ST_Point(1,2));
 st_astext
------------
 POINT(1 2)
```

□□□□□□□□□PostGIS□□□□□□□□□□□□Geometry□□□□□□□□□□□□□□□□□□□□□Geometry□□□□□□□□□

```
gis=# CREATE TABLE geo (
    geom GEOMETRY
);
gis=# INSERT INTO geo VALUES
    (ST_Point(1.0, 2.0)),
    ('LineString(0 0,1 1,2 1,2 3)'),
    ('Polygon((0 0, 1 0, 1 1,0 1,0 0))'),
    ('MultiPoint(1 2,3 4)');
```

18.3.3 [illegible]

[illegible]WKT[illegible]GeoJSON[illegible]

```
gis=# SELECT
    ST_AsText(geom)    AS wkt,
    ST_AsGeoJSON(geom) AS json
FROM geo;
```

[illegible]18-1[illegible]

[illegible]18-1 [illegible]

wkt	json
POINT(1 2)	{"type":"Point","coordinates":[1,2]}
LINESTRING(0 0,1 1,2 1,2 3)	{"type":"LineString","coordinates":[[0,0],[1,1],[2,1],[2,3]]}
POLYGON((0 0,1 0,1 1,0 1,0 0))	{"type":"Polygon","coordinates":[[[0,0],[1,0],[1,1],[0,1],[0,0]]]}
MULTIPOINT(1 2,3 4)	{"type":"MultiPoint","coordinates":[[1,2],[3,4]]}

[illegible]

```
geo=# SELECT
ST_AsLatLonText(ST_Point(116.321367,39.966956));
        st_aslatlontext
--------------------------------
    39°58'1.042"N 116°19'16.921"E
```

18.3.4 [illegible]

PostGIS[illegible]SQL[illegible]

[illegible]

```
geo=# SELECT ST_Point(1,1) <-> ST_Point(2,2);
1.4142135623730951
```

[illegible]1[illegible]1[illegible]2[illegible]2[illegible]

[illegible]A[illegible]116.321367[illegible]39.966956[illegible]B[illegible]116.315346[illegible]39.997398[illegible]

```
gis=# SELECT ST_Point(116.321367, 39.966956) <->
ST_Point(116.315346, 39.997398);
-- 0.0310317225933419[illegible]
```

□□□□□□□□□□□□□4326□□□□□□□□WGS84□□□□GPS□□□□□□□□□□□□□□□□□□□□□□□□□3.4km□□□□□□□

```
gis=# SELECT ST_AsText(ST_GeomFromText('POINT(116.321367
39.966956)', 4326)) :: GEOGRAPHY <->
    ST_AsText(ST_GeomFromText('POINT(116.315346 39.997398)',
4326)) :: GEOGRAPHY;
3423.653480690467
```

□□□□□□□□□□□□R□□□□□□□□□□ST_length□□WGS84□□□□□□□□□□□□□□□□□

```
-- □R□□□□MultiLineString□□
gis=# SELECT
ST_length(ST_GeomFromText('MULTILINESTRING((116.351494
39.976407,116.353159 39.976395,116.353365 39.976479),
(116.350712 39.976471,116.350984 39.976406,116.351494
39.976407,116.351755 39.976479),(116.351516
39.976558,116.346836 39.976502),(116.350984
39.976406,116.345421 39.976355,116.344646
39.976304,116.343376 39.976326,116.343353
39.976185,116.343285 39.976147,116.340918 39.976076),
(116.332306 39.976515,116.326279 39.976308),(116.354314
39.976488,116.347592 39.976436,116.347855 39.97637),
(116.354302 39.976574,116.351516 39.976558,116.351407
39.976616,116.349808 39.976603,116.346841
39.976569,116.346836 39.976502,116.34048
39.976443,116.335918 39.976538),(116.343194
39.976399,116.340565 39.976374,116.333921
39.976471,116.331997 39.976417,116.332003 39.976342),
(116.346841 39.976569,116.344304 39.97658,116.34256
39.976526,116.336051 39.976581,116.335918
39.976538,116.333922 39.976571,116.332306
39.976515,116.332108 39.976573,116.326431
39.976373,116.326018 39.976391),(116.331997
39.976417,116.323548 39.976145),(116.347592
```

```
39.976436,116.343194 39.976399,116.343376
39.976326,116.340209 39.976302,116.336219
39.976364,116.335423 39.976342,116.333617
39.976397,116.327658 39.976201,116.326968
39.976147,116.325406 39.976134,116.324303
39.97606,116.323696 39.976076),(116.317584
39.97613,116.320373 39.976186,116.323807
39.976325,116.324613 39.976315,116.325891
39.976391,116.326018 39.976391,116.326279
39.976308,116.317613 39.976048),(116.31983
39.975958,116.32005 39.976033,116.317835
39.975977,116.317637 39.975969,116.317777
39.97589,116.323696 39.976076,116.323548 39.976145,116.32005
39.976033))',4326)::geography);
12314.809569165007
```

[illegible]R[illegible]12[illegible]

[illegible]ST_Area[illegible]ST_Polygon[illegible]

```
gis=# SELECT ST_Area(ST_GeomFromText('POLYGON((116.402408
39.915326,116.402822 39.91533,116.402847
39.915036,116.402434 39.915035,116.402426
39.91495,116.402474 39.91397,116.402493 39.91367,116.402505
39.913385,116.40251 39.913269,116.401675 39.91323,116.400835
39.913238,116.400346 39.913231,116.398979
39.913198,116.398359 39.913159,116.398337
39.913153,116.398352 39.912786,116.398317
39.912738,116.398274 39.9127,116.398221 39.912661,116.39806
39.912645,116.398029 39.912642,116.39792
39.912639,116.397909 39.912638,116.397904 39.912668,116.3971
39.912644,116.396687 39.91261,116.396659
39.912609,116.396553 39.912606,116.39653
39.912605,116.396519 39.912746,116.396401
39.912747,116.396354 39.912759,116.396331 39.91294,116.39632
39.913079,116.396302 39.913093,116.396264
39.913105,116.396167 39.913097,116.394376
39.91302,116.393154 39.912964,116.392198
39.912919,116.392031 39.912918,116.392004
```

```
39.913377,116.391997 39.913535,116.391984
39.91369,116.391974 39.91388,116.391967 39.914106,116.391962
39.914287,116.391948 39.914579,116.391938
39.914699,116.391921 39.915047,116.3919 39.915444,116.391854
39.916108,116.391789 39.917649,116.391766
39.918064,116.391707 39.919443,116.391637
39.920661,116.391595 39.921275,116.391579
39.921716,116.391521 39.922832,116.391519
39.922907,116.39159 39.922907,116.392208
39.922931,116.394282 39.923007,116.396813
39.923109,116.398056 39.923155,116.40073
39.923241,116.400982 39.923244,116.401861
39.923272,116.401931 39.923274,116.402009
39.923278,116.402013 39.923148,116.402028
39.922837,116.402043 39.922627,116.402053
39.922343,116.402065 39.922116,116.402075
39.921807,116.40208 39.921672,116.402115
39.921004,116.402133 39.920698,116.402216
39.918928,116.402229 39.918673,116.402269
39.91807,116.402291 39.917566,116.402317
39.917029,116.402343 39.916641,116.402363
39.916035,116.402366 39.915835,116.402408 39.915326))',
4326)::GEOGRAPHY);
1005170
```

□□□□□□□□□□□□□□□□□□□□□□□□□□□□

18.4 [illegible]

[illegible]LBS[illegible]1[illegible]

[illegible]A[illegible]116.321367[illegible]39.966956[illegible]1[illegible]111[illegible]

```
gis=# CREATE TABLE stations(
    name       TEXT,
    longitude DOUBLE PRICISION,
    latitude  DOUBLE PRICISION,
);
-- [illegible]
gis=# SELECT name FROM stations WHERE longitude BETWEEN
116.312358 AND 116.330376
AND latitude BETWEEN 39.957947 AND 39.975965;
```

[illegible]GeoHash[illegible]

[illegible]PostGIS[illegible]PostGIS[illegible]
SQL[illegible]

```
gis=# SELECT
    name,
    ST_Point(116.321367, 39.966956) :: GEOGRAPHY <->
position :: GEOGRAPHY AS distance
FROM stations
WHERE ST_Point(116.321367, 39.966956) :: GEOGRAPHY <->
position :: GEOGRAPHY < 500
ORDER BY ST_Point(116.321367, 39.966956) :: GEOGRAPHY <->
position :: GEOGRAPHY;
```

18.4.1 □□□□

□100□□1□□□□□POI□□□□□□□A□□□□□□1000□POI□□□□□□□

```
gis=# SELECT name FROM poi ORDER BY position <->
ST_Point(116.458855, 39.909863) LIMIT 1000;
```

□□□□□□□3□□□□□□□PostgreSQL□□□GIST□□□□□□□□

```
gis=# CREATE INDEX CONCURRENTLY idx_poi_position_gist ON poi
USING gist (position);
```

□□□□□□□□□□□□□□□

```
                                QUERY PLAN
------------------------------------------------------------
```

```
-----------------
Limit  (cost=0.42..9.73 rows=10 width=31)
    ->  Index Scan using idx_poi_position_gist on poi
(cost=0.42..58440964.86 rows=62750132 width=31)
        Order By: ("position" <->
'0101000000CAA65CE15D1D5D40946B0A6476F44340'::geometry)
```

□□□□□□□1ms□□□□□□□□□□□□□□□□

```
geo=# SELECT name FROM poi ORDER BY position <->
ST_Point(116.458855, 39.909863) LIMIT 10;
             name
------------------------------
...
...
...
(10 rows)
Time: 0.993 ms
```

18.4.2 □□□□

□□

□□□

```
-- □□□□ (AOI, Area of Interest)
gis=# CREATE TABLE aoi(
    name    TEXT,
    bound   GEOMETRY
```

```
)
-- □□A□□□□□□□□□□□
gis=# SELECT name FROM aoi WHERE ST_Contains(bound,
ST_Point(116.458855, 39.909863));
...
□□□□□
□□□□
□□□□□
...
```

18.5 [illegible]

PostGIS[illegible]PostGIS[illegible]PostGIS[illegible]

[illegible]PostGIS[illegible]PostGIS[illegible]PostGIS[illegible]